Report

DFG reference number and title

SCHU 1566/3-1 – Neuroframes – Neuronale Grundlagen von Funktionalbegriffen

Personnel

Project head:

- Univ.-Prof. Dr. Gerhard Schurz

Scientific Personnel

- Dr. Markus Werning (BAT IIa/TV-Ü13, 1.8.2005–31.7.2008)

Student research assistants (ten hours/week)

- Jens Fleischhauer (1.10.2005–28.2.2007)
- Hakan Beseoglu (1.1.2006–31.7.2008)
- Patrick Albertini (01.03.2007–30.06.2007)

1 State of knowledge and goals

Frame theory provides us with a universal account not only for categorization and its link to action-control, but also for the decomposition of concepts (Barsalou, 1992). Frames are recursive attribute-value structures. Attributes assign unique values to objects and thus describe functional relations. The values can be structured frames themselves. A frame is defined for a large domain of things and contains a fixed set of attributes (e.g., color, form), each of which allows for a number of different values (red, green, ... ; round, square, ...). The attributes in question are not constrained to perceptual modalities, but may as well involve attributes of motor affordances. Frames can be nested hierarchically and mutual constraints between attributes, for instance between states of an object and actions directed to it and between larger frames, can be incorporated (see Figure 1).

The main goal of project B2 is to develop a biologically plausible and philosophically motivated model for the cortical realization of frames and functional concepts they contain, so that frames may also be regarded as a plausible model for conceptual decomposition from a neurobiologically informed point of view.

The following sub-goals were addressed during the three year period: (i) Frames can be justified as an adequate format for lexical decomposition. (ii) Frames can be implemented by oscillatory neural networks. (iii) A hierarchical binding mechanism needs to be postulated to achieve the neural representation of part-whole relations.

2 Results and their significance

In our context, the most important distinction in the domain of concepts is the one between attributive concepts and substance concepts. Attributive concepts represent features of objects that are volatile in the sense that one and the same object can be subsumed under different attributive concepts at different times: An object may change its color, size, or speed, but still continue to exist. [Blue] is a paradigmatic attributive concept and thus serves as a value for the attribute color. Substance concepts, in contrast, are governed by the identity conditions of objects: a mug ceases to exist when it no longer falls under the substance concept [mug], say, because it has been shattered. Substance concepts serve to re-identify things over time, in spite of their contingent changes of attributes, and allow us to gather, store and update information in a systematic and enduring way (Millikan, 1998). They are typically expressed by concrete nouns — for example, in English by names of individuals like *mama*, names of kinds like *mouse* and names of stuffs like *milk*. Attributive concepts, by contrast, are typically expressed in English by adjectives or abstract nouns: *blue*(-*ness*), *warm*(-*th*), *lucid*(-*ity*).

So far, the project has been mainly concerned with the decomposition of substance concepts by means of functional concepts with attributive concepts as their values. The perspective developed in the project has resulted in the theory of neuro-frames (Werning & Maye, 2007; Petersen & Werning, 2007). The theory of neuro-frames holds that (i) substance concepts are decomposable into less complex concepts, that (ii) the decompositional structure of a substance concept can be rendered by a recursive attribute-value structure or frame, that (iii) the neural realization of a substance concept is distributed over assemblies of neurons and meta-assemblies thereof, that (iv) those neurons pertain to neural maps for various attributes in many afferent and efferent regions of cortex and that (v) an appropriate mechanism of binding the distributed information together into the neural realization of the substance concept is the mechanism of neural synchronization.

One can anatomically identify cortical correlates for many attributes involved in perceptual processing. These areas often exhibit a twofold topological structure and justify the notion of a feature map: (i) a receptor topology (e.g., retinotopy in vision, somatotopy in touch): neighboring regions of neurons code for neighboring regions of the receptor; and (ii) a feature topology: neighboring regions of neurons code for similar features. In connection with the monkey, more than 30 cortical areas forming feature maps have been experimentally identified for vision alone (Felleman & van Essen, 1991).

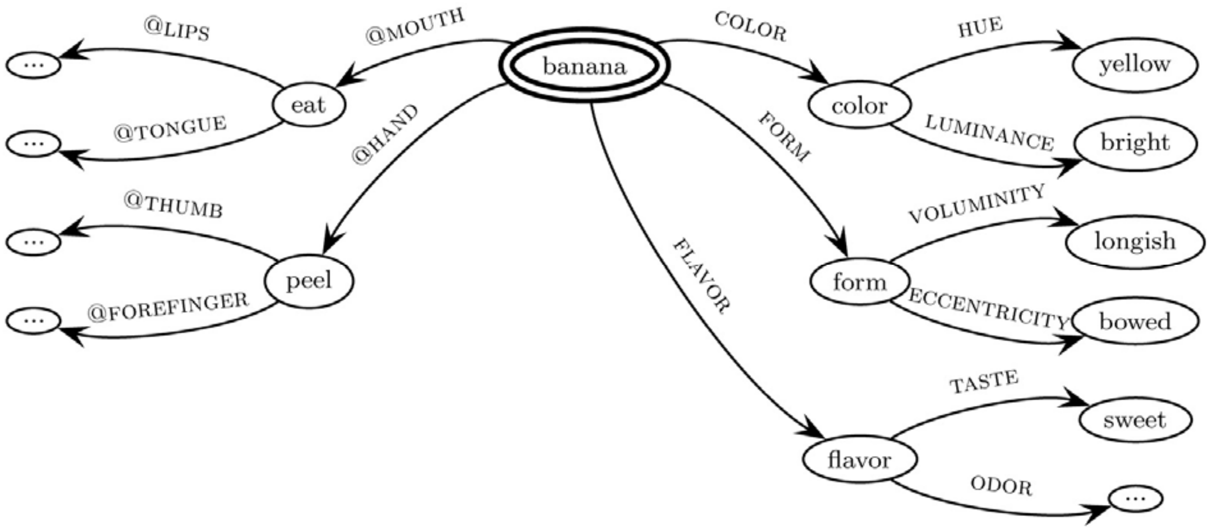
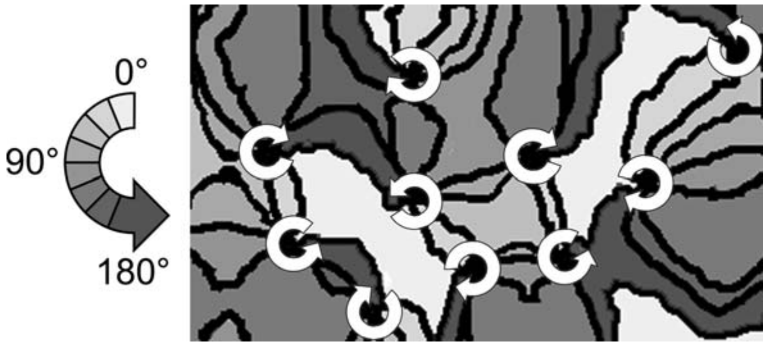


Figure 1. Hypothetical fragment of the frame for the substance concept [banana]. The substance concept to be decomposed is marked by a double-circle as the referring node of the frame. The labeled arrows denote attributes, the nodes their values. Nodes are regarded as concepts and thus as conceptual parts of the central concept. In English, feature attributes (shown on the right) are frequently lexicalized; their arguments typically enter possessive constructions like *The color of the banana is yellow* or *The banana has the color yellow*. Based on linguistic and neurobiological evidence, we assume that affordances often relate to body parts and have adopted the convention "@ + body part". Formally, attributes are mappings from domains of some type into domains of some other type. Petersen & Werning (2007) provide an explicit account of frames using a calculus of typed feature hierarchies and incorporating typicality effects.

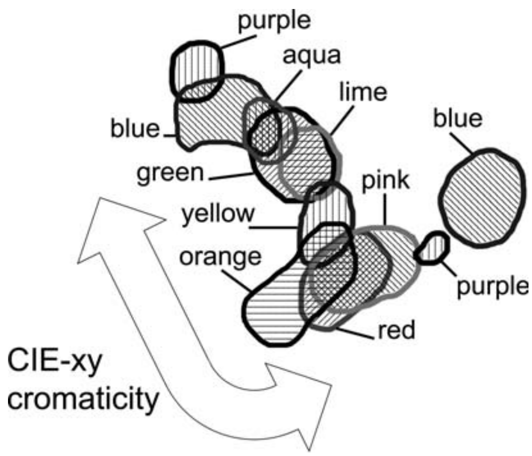
Affordance attributes, too, may be parts of frames and seem to have cortical correlates, predominantly in the premotor cortex (Werning, 2008). The cortical organization of motor control with regard to the effectors follows topological principles similar to the cortical organization in perception with respect to the receptors. The discovery of the so-called mirror neuron system (Rizzolatti & Craighero, 2004, for review) may provide a basis for integrating affordances into frames. Figure 2 depicts a number of neural maps that relate to attributes of frames.

The fact that values of different attributes may be instantiated by the same object but are processed in distinct regions of the cortex poses the problem of how this information is integrated in an object-specific fashion: the binding problem (Treisman, 1996). How can color and taste of a banana be represented in distinct regions of the cortex, although they are still part of the representation of one and the same object?

a)



b)



c)

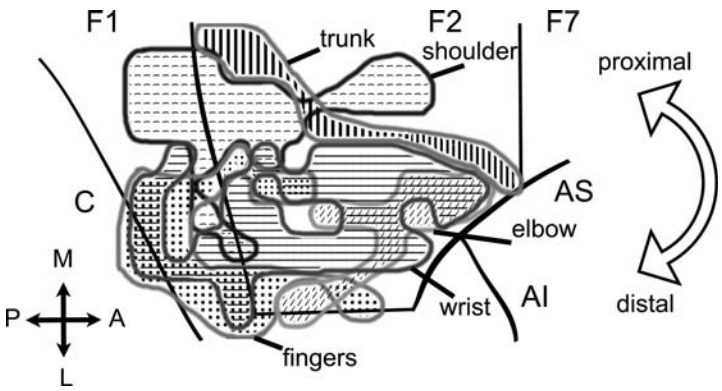


Figure 2. Cortical realizations of frame attributes. a) Fragment (ca. 4mm²) of the neural feature map for the attribute orientation of cat V1 (adapted from Crair et. al., 1997). The arrows indicate the polar topology of the orientation values represented within each hypercolumn. Hypercolumns are arranged in a retinotopic topology. b) Color band (ca. 1 mm²) from the thin stripes of macaque V2 (adapted from Xiao et. al., 2003). The values of the attribute color are arranged in a topology that follows the similarity of hue as defined by the Commission Internationale de l'Eclairages (xy-cromaticity). The topology among the various color bands of V2 is retinotopic. c) Neural map (ca. 250 mm²) of forelimb movement in macaque primary motor (F1) and dorsal premotor cortex (F2, F7) (adapted from Raos et al., 2003). The overarching topology is somatotopic from proximal to distal movement as shown by the arrow. Due to the size of the region one might expect it to comprise maps for more specific motor attributes. C: central sulcus, AS and AI: superior, respectively, inferior arcuate sulcus.

A prominent and experimentally well supported solution postulates oscillatory neural synchronization as a mechanism for binding: Clusters of neurons that are indicative for different properties sometimes show synchronous oscillatory activity, but only when the properties indicated are instantiated by the same object in the perceptual field, otherwise they fire asynchronously. Thus, synchronous oscillation might be regarded as fulfilling the task of binding various property representations together to form the representation of an object having these properties (Singer, 1999). Using oscillatory networks as biologically motivated models it could be demonstrated by mechanisms of synchronization how the topological organization of information in the cortex could yield a logically structured semantics of concepts (Maye & Werning, 2007, see Figure 3). Compositionality theorems have been provided (Werning, 2005d). Oscillation functions play the role of object concepts. Clusters of feature sensitive neurons assume the role of attributive concepts. Schnitzler et al. (2006) could experimentally demonstrate the essential role of neural synchronization for action control. This might justify the extension of the synchrony-based neuro-frame approach from features to affordances.

Support for the theory of neuro-frames also comes from a number of neuro-linguistic studies. Based on a review of neurobiological data, Pulvermüller (1999) suggests that neural assemblies that pertain to the sensori-motor cortices and are bound by neural synchronization play an important role in understanding the meanings of words. FMRI studies (Pulvermüller, 2005) regarding the understanding of verbs allude to a differential top-down activation of motor and pre-motor areas. We know that the understanding of concrete nouns like *hammer*, for which not only features but also affordances are salient, results in an activity distributed over the premotor and the visual cortex (Martin et. al. 1996). The hypothesis that words for substance concepts arouse more widely distributed activity than words for attributive concepts has also been confirmed by EEG studies (Rappelsberger et al., 2000).

The neuro-semantic approach used here is based on simulatory studies on perceptual oscillatory networks by Markus Werning and his co-operators (Maye & Werning, 2007; Maye, A., Werning, M., König, P., & Engel, A., 2005). By means of an eigenmode analysis it could be shown that the mechanism of oscillatory neural synchronization may subserve the realization of the semantics of a monadic first order predicative language (Werning, 2005a). The virtues of classicist and connectionist approaches could be attained, while the shortcomings of both paradigms were avoided Extensive work has been done regarding the issue of compositionality (Werning, Hinzen, & Machery (Eds., 2009); Werning, Machery, & Schurz, (Eds., 2005); Machery, Werning, & Schurz, (Eds., 2005); Machery, Schurz, & Werning (2005). The reasons for compositionality (Werning, 2005e; Werning et al. 2009) and the prospects of implementation have been discussed (Werning & Maye, 2005; Werning, 2009). The question of compositionality for prototype concepts as represented by prototype frames has been investigated by Schurz (2009). The latter will show that prototype attributes satisfy non-monotonic (default) inheritance relations, which can be modelled by the basic rules of probability logic and can be explained by their function as speed and efficient (though by no means infallible) tools of diagnostic reasoning in evolution as cognition. The evolutionary justification of probabilistic default reasoning patterns in human cognition has already been formulated in Schurz (2007).

Ambiguous und illusionary object perceptions like the Kanizsa-illusion (Werning & Maye, 2006) could be integrated in the approach as well as part-whole relations (Werning &

Maye, 2005). Initial attempts to integrate recursive embedding into neuro-frames by means of coherency chains and hierarchical binding were successful. From a philosophical point of view the topic of conscious perception and the debate between pictorial and conceptual representations has been dealt with (Werning, 2005c). Issues of higher-order intentionality (Abraham, Werning, Rakoczy, von Cramon, & Schubotz, 2008), metaphorization (Werning, Fleischhauer, & Beseoglu, 2006) as well as evolutionary and developmental issues (Werning, 2005b, 2008) have also been approached.

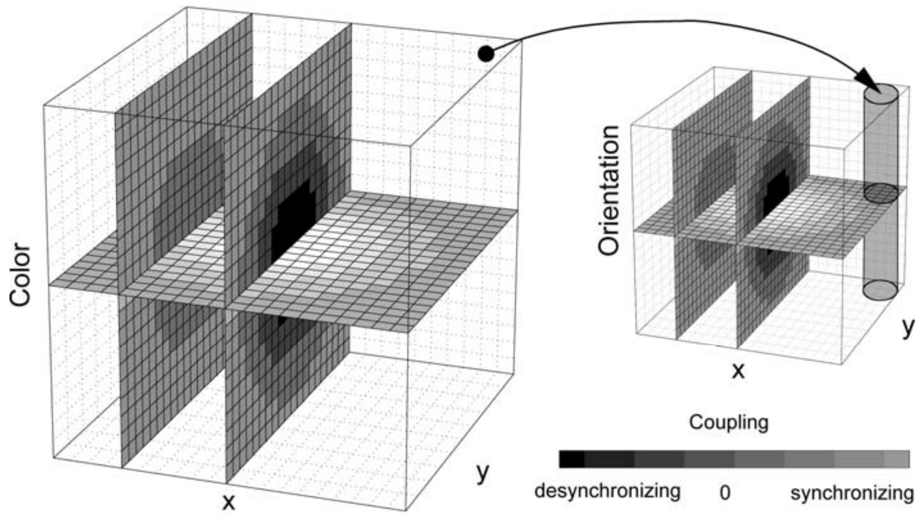


Figure 3. Oscillatory network. The network topology reflects the receptor topology (xy-plane) and the feature topology (z-axis) of the neural maps. Each module realizes one attribute. The layers in each module realize the attribute values. Oscillators activated by neighboring stimulus elements with like attribute values synchronize (light gray). Oscillators activated by neighboring stimulus elements with unlike attribute values de-synchronize (dark gray). The layers of different modules are connected in a synchronizing way that respects the common receptor topology. (From Maye & Werning, 2007).

3 Relation of work schedule to outcome

i) Lexical decomposition by means of frames	months 1-12	Goal achieved. A formal theory of lexical decomposition by means of frames has been worked out and published. The integration of affordance attributes is still virulent (see follow-up application).
ii) Implementation of frames in oscillatory networks	months 13-24	Goal achieved (for simple frames). Numerous publications.
iii) Hierarchical binding and the neuronal representation of part-whole relations	months 15-36	Goal partially achieved. Mechanism of hierarchical binding have been suggested, implemented and the results have been published.

iv) The analysis and implementation of constraints	months 37-60	Future goal.
v) The analysis and implementation of abstract functional concepts	months 61-72	Future goal. Preliminary work on abstraction by metaphorization has been carried out and the results have been published.

4 Cooperations

4.1 within the research unit

- A1 Exchange over the typology of functional concepts and issues of compositionality. See Löbner (forthcoming) in Werning, M., Hinzen, W., & Machery E. (Eds., 2009). Exchange over grammaticalizations in non-European languages with regard to the expression of affordances. See Werning (2008).
- A3 Exchange over issues of compositionality in formal semantics. See Zimmermann (forthcoming) in Werning, M., Hinzen, W., & Machery E. (Eds., 2009)
- A4 Exchange over the linguistic means to express affordances. Exchange with regard to the acquisition of nouns and verbs in development. See Werning (2008).
- B1 Development of a formal theory of frames as a means for conceptual decomposition (see Petersen & Werning, 2007)
- B3 Exchange of the notion of substance concepts in the history of philosophy and contemporary metaphysics.

4.2 with external partners

4.2.1 national

- Prof. Dr. Andreas Engel and Dr. Alexander Maye, Institute for Neurophysiology, University Hospital Hamburg-Eppendorf
- Prof Dr. Peter König and Dr. Sabine Weiss, Institute for Cognitive Science, University of Osnabrück
- PD Dr. Ricarda Schubotz and Dr. Anna Abraham, MPI for Human Cognitive and Brain Sciences, Leipzig
- Dr. Hannes Rackocy, MPI for Evolutionary Anthropology, Leipzig

4.2.1 international

- Prof. Dr. Wolfram Hinzen, Department of Philosophy, University of Durham
- Dr. Edouard Machery, Department of History and Philosophy of Science, University of Pittsburgh
- Prof. Dr. Gernot Kleiter, Prof. Dr. Heinz Wimmer, Prof. Dr. Joseph Perner, Institute for Psychology, Universität Salzburg

- Prof. Dr. Jeff Pelletier, Canada Chair of Cognitive Science, Simon Fraser University, Vancouver
- Prof Dr. Brian McLaughlin and Michaela Tacca, MA, Department of Philosophy, Rutgers University, New Jersey
- Prof. Dr. Jim Hurford and Dr. Simon Kirby, School of Philosophy, Psychology and Linguistics, University of Edinburgh
- Prof. Dr. Peter Gärdenfors, Department of Cognitive Science, Lund University
- Prof. Dr. Ruth Millikan, Department of Philosophy, University of Connecticut
- Prof. Dr. Terry Horgan, Department of Philosophy, University of Arizona at Tuscon
- Prof. Dr. Peter Hagoort, MPI for Psycholinguistics and

5 List of publications in the project period beginning 2005

I. Refereed Publications

a) in Journals

Abraham, A., Werning, M., Rakoczy, H., von Cramon, D. Y., & Schubotz, R. I. (2008). Minds, persons, and space: An fMRI investigation into the relational complexity of higher-order intentionality. *Consciousness and Cognition.* (In press)

Werning, M. (2008). The ‘complex first’ paradox: Why do semantically thick concepts so early lexicalize as nouns? *Interaction Studies, 9*(1), 66-82. (In press)

Maye, A., & Werning, M. (2007). Neuronal synchronization: From dynamic feature binding to object representations. *Chaos and Complexity Letters, 2*(2/3), 315-325.

Werning, M., & Maye, A. (2007). The cortical implementation of complex attribute and substance concepts: Synchrony, frames, and hierarchical binding. *Chaos and Complexity Letters, 2*(2/3), 435-452.

Werning, M. (2005a). The temporal dimension of thought: Cortical foundations of predicative representation. *Synthese*, *146*(1/2), 203-224.

Werning, M. (2005b). The ‘complex first’ paradox: What’s needed to compose neurally distributed lexical meanings? *Behavioral and Brain Sciences*, *28.* (Online suppl.).

b) in Proceedings

Schurz, G. (2007). Human conditional reasoning explained by non-monotonicity and probability: An evolutionary account. In S. Vosniadou, D. Kayser, & A. Protopapas (Eds.), *Proceedings of the 2nd European Cognitive Science Conference* (pp. 628-633). New York: Lawrence Erlbaum Associates.

Werning, M., & Maye, A. (2006). The neural basis of the object concept in ambiguous and illusionary perception. In R. Sun, & N. Miyake (Eds.), *Proceedings of the Twenty-eighth Annual Conference of the Cognitive Science Society* (pp. 876-881). London: Lawrence Erlbaum Associates.

Werning, M., Fleischhauer, J., & Beseoglu, H. (2006). The cognitive accessibility of synaesthetic metaphors. In R. Sun, & N. Miyake (Eds.), *Proceedings of the Twenty-eighth Annual Conference of the Cognitive Science Society* (pp. 2365-2370). London: Lawrence Erlbaum Associates.

Werning, M. (2005c). Time structure and the structure of perception. In F. Stadler, & M. Stöltzner (Eds.), *Time and history: Contributions of the Austrian Ludwig Wittgenstein Society* (pp. 336-339). Kirchberg am Wechsel: Austrian Ludwig Wittgenstein Society.

Werning, M., & Maye, A. (2005). Frames, coherency chains and hierarchical binding: The cortical implementation of complex concepts. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the Twenty-seventh Annual Conference of the Cognitive Science Society* (pp. 2347-2352). London: Lawrence Erlbaum Associates.

Maye, A., Werning, M., König, P., & Engel, A. (2005). Advancing dynamic binding theory: Implementation of complex concepts. In H. Zimmermann, & K. Krieglstein (Eds.), *Proceedings of the 6th Meeting of the German Neuroscience Society / 30th Göttingen Neurobiology Conference* (p. 36B). Heidelberg: Elsevier.

c) in Anthologies

Petersen, W., & Werning, M. (2007). Conceptual fingerprints: Lexical decomposition by means of frames – a neuro-cognitive model. In U. Priss, S. Polovina, & R. Hill (Eds.), *Conceptual structures: Knowledge architectures for smart applications* (LNAI 4604) (pp. 415-428). Heidelberg: Springer-Verlag.

Schurz, G. (to appear 2009). Prototypes and their composition from an evolutionary point of view. In M. Werning, W. Hinzen, & E. Machery (Eds.), *The Oxford handbook of compositionality*. Oxford: Oxford University Press.

Werning, M. (to appear 2009). The neural foundation of semantic composition. In M. Werning, W. Hinzen, & E. Machery (Eds.), *The Oxford handbook of compositionality*. Oxford: Oxford University Press.

II. Non-refereed Publications

a) in Journals

b) in Proceedings

c) in Anthologies

Werning, M. (2005d). Neuronal synchronization, covariation, and compositional representation. In E. Machery, M. Werning, & G. Schurz (Eds.), *The compositionality of meaning and content* (Vol. II: Applications to Linguistics, Philosophy and Neuroscience, pp. 283-312). Frankfurt: Ontos Verlag.

Werning, M. (2005e). Right and wrong reasons for compositionality. In M. Werning, E. Machery, & G. Schurz (Eds.), *The compositionality of meaning and content* (Vol. I: Foundational Issues, pp. 285-309). Frankfurt: Ontos Verlag.

Machery, E., Schurz, G., & Werning, M. (2005). Preface. In M. Werning, E. Machery, & G. Schurz (Eds.), *The compositionality of meaning and content* (Vol. I: Foundational Issues, pp. 7-21). Frankfurt: Ontos Verlag.

Werning, M., Hinzen, W., & Machery, E. (to appear 2009). Compositionality: Reasons, problems and perspectives. In M. Werning, W. Hinzen, & E. Machery (Eds.), *The Oxford handbook of compositionality*. Oxford: Oxford University Press.

III. Monographs und Editions

Werning, M., Hinzen, W., & Machery, M. (Eds.) (to appear 2009). *The Oxford Handbook of Compositionality*. Oxford: Oxford University Press (contracted, ca. 250.000 words).

Werning, M., Machery, E., & Schurz, G. (Eds.) (2005). *The compositionality of meaning and content* (Vol. I: Foundational Issues). Frankfurt: Ontos Verlag.

Machery, E., Werning, M., & Schurz, G. (Eds.) (2005). *The compositionality of meaning and content* (Vol. II: Applications to Linguistics, Psychology and Neuroscience). Frankfurt: Ontos Verlag.

6 Activities

Guests invited to Düsseldorf by project B2:

- Dr. Sabine Weiss (Univ. Osnabrück, Germany)

- Dr. Marcel Bastiaansen (MPI Nijmegen, Holland)
- Prof. Dr. Friedemann Pulvermüller (Univ. Cambridge, UK)
- Prof. Dr. Peter Gärdenfors (Univ. Lund, Sweden)
- Prof. Dr. Xiang Chen (California Lutheran University, USA)

Conference organization

Markus Werning was co-organizer of the conference *Concept Types and Frames in Language, Cognition and Science* (Düsseldorf, 20-22 August 2007)

Editorial activities

Markus Werning is editor of the contracted *Oxford Handbook of Compositionality* (250,000 words, Oxford University Press, to appear 2009)

References

Arbib, M., Conklin, E., & Hill, J. (1987). *From schema theory to language*. Oxford: OUP.

Barrett, H., & Kurzban, R. (2006). Modularity in cognition. *Psychological Review*, *113*, 628-647.

Barsalou, L. (1992). Frames, concepts, and conceptual fields. In A. Lehrer, & E. Kittay (Eds.), *Frames, fields, and contrasts* (pp. 21-74). Hillsdale, NJ: Erlbaum.

Barsalou, L. (2005). Situated conceptualization. In H. Cohen, & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 619-650). St. Louis: Elsevier.

Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence*, *47*, 139-159.

Clifton, C., & Ferreira, F. (1987). Modularity in sentence comprehension. In J. Garfield (Ed.), *Modularity in knowledge representation and natural-language understanding* (pp. 277-290). MIT Press.

Crair, M. C., Ruthazer, E. S., Gillespie, D. C., & Stryker, M. P. (1997). Ocular dominance peaks at pinwheel center singularities of the orientation map in cat visual cortex. *Journal of Neurophysiology*, *77*(6), 3381-3385.

Felleman, D., & van Essen, D. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*, 1-47.

Fodor, J. (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.

Fodor, J. (1998). *Concepts: Where Cognitive Science Went Wrong*. Oxford: Oxford University Press.

Friederici, A. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*, 78-84.

Keijzer, F. (2002). Representation in dynamical and embodied cognition. *Cognitive Systems Research*, *3*, 275-288.

Martin, A., Wiggs, C. L., Ungerleider, L., & Haxby, J. V. (1996). Neural correlates of category specific knowledge. *Nature*, *379*, 649-652.

Millikan, R. G. (1998). A common structure for concepts of individuals, stuffs and real kinds: More mama, more milk, and more mouse. *Behavioral and Brain Sciences*, *21*, 55-100.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences*, *22*, 253-279.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, *6*, 576-582.

Raos, V., Franchi, G., Gallese, V., & Fogassi, L. (2003). Somatotopic organization of the lateral part of area F2 (dorsal premotor cortex) of the macaque monkey. *Journal of Neurophysiology*, *89*(3), 1503-1518.

Rappelsberger, P., Weiss, S., & Schack, B. (2000). Coherence and phase relations between EEG traces recorded from different locations. In R. Müller (Ed.), *Time and the brain* (pp. 297-330). Harwood.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169-192.

Schnitzler, A., Timmermann, L., & Gross, J. (2006). Physiological and pathological oscillatory networks in the human motor system. *Journal of Physiology*, *99*, 3-7.

Singer, W. (1999). Neuronal synchrony: A versatile code for the definition of relations? *Neuron*, *24*, 49-65.

Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology, 6*, 171-178.

Xiao, Y., Wang, Y., & Felleman, D. (2003). A spatially organized representation of colour in macaque cortical area V2. *Nature*, *421*, 535-539.