1. Logic, Neural Networks, and Optimality Theory

(preparatory grant proposal)

2. Abstract of the problem statement and project goal

Ever since the discovery of neural networks, there has been a controversy between two modes of information processing. On the one hand symbolic systems have proven as indispensable for our understanding of higher intelligence, essentially when cognitive domains like language and reasoning are examined. On the other hand we believe that intelligence resides in the brain, where computation appears to be numerical, not symbolic; parallel, not serial; rather distributed, not as highly localized as in symbolic systems. We claim that this controversy can be resolved by a unified theory of cognition only – one that integrate both aspects of cognition and assigns the proper roles to symbolic computation and numerical neural computation.

The overall goal in this project is to develop and study formal systems which are suitable to ground the formal basis for such a unified theory. The proposal suggests that the instruments of modern logic and model theoretic semantics are appropriate for analyzing certain aspects of dynamical systems such as inferring and learning in neural networks. We suggest that an active dialogue between the traditional symbolic approaches to logic, information and language and the connectionist paradigm is possible and fruitful. An essential component of this dialogue refers to *Optimality Theory* – taken as a theory that likewise aims to overcome the gap between symbolic and neuronal systems.

3. Applicant

PD Dr. Reinhard Blutner

4. Institutional environment

ILLC, University of Amsterdam

5. Composition of the research team

name	Relevant expertise	affiliation
dr. Reinhart Blutner	OT, neural networks, evolutionary game theory	Philosophy, UvA
prof. Michiel van Lambalgen	logic, perception, connectionism	Philosophy, UvA
dr. Henk Zeevat	OT, computational aspects	Computational Linguistics,UvA
dr. Paul Boersma	stochastic OT, learning	Phonology, UvA
dr. Helen de Hoop	OT, learning, connectionism	Pioneer Project & NWO-Cognition Project, KUN
dr. Petra Hendriks	OT, pragmatics	Dutch and AI, RUG; NWO-Cognition Project
prof. Henrïette de Swart	OT, syntax	Linguistics, Utrecht University; NWO-Cognition Project
dr. Gerhard Jäger	Stochastic OT, learning, evolutionary game theory	SFB-Project "Bidirection", ZAS
prof. Joan Bresnan	syntax, stochastic OT, language change	Linguistics, Stanford University
prof. David Beaver	logic, semantics, OT	Linguistics, Stanford University
prof. Paul Smolensky	connectionism, OT, foundation	Cognitive Science, JHU

RUG = Groningen University
UvA = University of Amsterdam
KUN = Nijmegen University
ZAS = Zentrum für Allgemeine Sprachwissenschaft
JHU = Johns Hopkins University

6. Time-frame

The period for subsidy is June 2003 – October 2003

7. Description of proposed activity

a. Contribution to the goals of the Cognition Programme

There are three related main themes of the project proposal:

1. <u>The gap between connectionism and symbolism</u>. What are the central general principles of computation in connectionist – abstract neural – networks? How can these principles be reconciled with those of symbolic computation?

   Earlier work (Gärdenfors, Blutner) attempted to overcome the gap between symbolism and connectionism by viewing symbolism as a high-level description of the properties of (a class of) neural networks. Combining methods of algebraic semantics and nonmonotonic logic, the possibility of *integrating* both modes of computation was demonstrated and it was shown that certain activities of connectionist networks can be interpreted as *nonmonotonic inferences*. However, there are many limitations in this treatment which should be overcome by the proposed research activity: first, the restriction to a strictly localist conception of neuronal representations, second the restrictions to very simple connectionist networks, and third the exclusion of

synchronous activation patterns and binding. The development of a neurosemantics of binding (with potential applications in visual perception, language comprehension and logical reasoning) is perhaps the most innovative aspect of the proposed research.

2. <u>Optimality theory and connectionism</u>. Optimality Theory (OT) has proposed a new computational architecture for cognition which claims to integrate connectionist and symbolic computation. However, there are many assumptions in standard OT where the neuronal underpinning is rather unclear. For example, it is an open issue whether the hierarchical encoding of constraint strengths can be explained in a theoretically satisfying way. An interesting and new hypothesis is that the hierarchical encoding of constraint strengths is correlated with the effect of automaticity in psychological processes [Similar distinctions play a role within paradigms like ACT-R]. If this hypothesis can be substantiated, the differences between automatic and controlled psychological processes can be understood as an emergent effect of the underlying neural computations.

3. <u>The problem of gradedness, probabilistic grammars and bi-directional learning</u>. Standard OT predicts a dichotomy of grammaticality judgments. This is due to the assumed strict ranking of the constraints. In many cases it conflicts with the observed phenomena – the spectrum of attested cases includes phonology, syntax, semantics, and pragmatics. There exist several proposals of how to deal with this situation: e.g., Smolensky's Harmony theory, Boersma's stochastic OT, Antilla's and De Hoop's assumption of partial constraint orderings. The issue that has to be resolved concerns the descriptive and explanatory adequacy of these theories and their neurosemantic and psychological underpinning. The most important innovative aspect concerns the combination of stochastic OT with bi-directional learning. In particular, it is proposed to investigate Gerhard Jäger's recent suggestion on *evolutionary learning* using the framework of genetic algorithms and evolutionary game theory. The main idea of this exercise is to shed new light on the underpinning of weak and strong bidirection in Optimality Theory, and to pave the way for an true explanation of the linguistically so important phenomenon of constructional iconicity.

With regard to the main goals of the Cognition Programme, the following aspects are of relevance:

- <u>Strengthening the methodological foundations of the cognitive sciences, and contributing to their coherence</u>. The basic idea in this respect is to use the instrument of logic as an instrument for analyzing dynamical systems in neuroscience.

- <u>Mathematical and philosophical integration of cognitive frameworks and levels</u>. Using the instruments of logic and algebraic semantics, the main goal is to overcome the gap between connectionism and symbolism. Examples are a neurosemantics of binding and a logic for OT learning.

- <u>Strengthening cognitive science institutionally</u>. One of aims of this proposal for a preparatory grant is to establish and strengthen contacts with related research groups in the Netherlands (especially with Helen de Hoop's pioneer project "*Case cross-linguistically*" and the NWO-Cognition Project "*Conflicts in Interpretation"*), Germany (SFB project "Bi-directional OT" led by Gerhard Jäger and Manfred Krifka, Berlin) and the USA (Stanford University & Johns Hopkins University)

- <u>Supporting integrative cognitive science education programs</u>. The proposed three research lines suggest different spin-offs with regard to educational programmes such as ESSLLI, NASSLLI, and the new master courses in Cognitive Science, Logic and Artificial Intelligence. For example, the logical reconstruction of OT systems and the reconstruction of OT learning can well be integrated into courses about computational aspects of OT. The connection between connectionism and OT can form part of a basic course "philosophy of mind", and insights into Paul Boersma's stochastic OT may possibly be integrated into courses about *Reasoning with Uncertainty*.

b. Practical results

The result will be a completely worked-out proposal for an integrative multidisciplinary research project in the NWO Cognition Programme.

c. Work plan, timeline and budget

- May - September 2003: Research assistant prepares workshops and makes preparations for the complete project proposal (literature search, contacts with research partners)

- June 2003: Workshop in Berlin in cooperation with the SFB project "Bi-directional OT" (ZAS, Berlin)

- September 2003: Workshop in Amsterdam/Nijmegen

- September 2003: Finishing the preliminary proposal for the full research project.

item	Euro
Salary costs research assistant (0.4 fte for 5 months)	7000
Traveling costs Netherlands, Germany	1400
Traveling costs foreign guests	2500
Workshops, including accommodation and foot	3200
total	14100

In addition, there are the following financial contributions of the
Pioneer Project "*Case cross-linguistically*" (KUN): 800 Euro
SFB project "Bi-directional OT" (ZAS, Berlin): 800 Euro

8. Short curriculum vitae of the applicant

Reinhard Blutner is a lecturer in Artificial Intelligence and Cognitive Philosophy at the University of Amsterdam and a .*Privatdozent* at the Humboldt-University in Berlin. He began his scientific career in theoretical physics and shifted later to artificial intelligence, cognitive psychology and theoretical linguistics. In his work he integrates insights from connectionist psychology, logics, computer science and cognitive linguistics.

9. Five key publications of the applicant

Blutner, R. (1996). 'Normality' in update semantics. In T. Galloway & M. Simons (Eds.), *SALT 5 Proceedings 1995* (pp. 19- 36). Cornell University: Cornell Linguistic Publications.
Blutner, R. (1997). Nonmonotonic logic and neural networks. In P. Dekker & M. Stokhof & Y. Venema (Eds.), *Proceedings of the eleventh Amsterdam Colloquium* (pp. 79-84): ILLC/Department of Philosophy, University of Amsterdam.
Blutner, R. (1998). Lexical pragmatics. *Journal of Semantics, 15*, 115-162.
Blutner, R. (2000). Some aspects of optimality in natural language interpretation. *Journal of Semantics, 17*, 189-216.
Blutner, R., & Zeevat, H. (Eds.). (in print). *Optimality Theory and Pragmatics*: Palgrave Macmillan.