# Grammar Specification

Categorial approaches to polymorphism and discontinuity


Michael Moortgat

(editor)


DYANA-2


Dynamic Interpretation of Natural Language

ESPRIT Basic Research Project 6852

Deliverable R1.3.A

August 1993

# Grammar Specification

Categorial approaches to polymorphism and discontinuity

Michael Moortgat
Rijksuniversiteit Utrecht
(editor)

authors:
Martin Emms
Johan van Benthem
Glyn Morrill

DYANA-2

Dynamic Interpretation of Natural Language
ESPRIT Basic Research Project 6852

Deliverable R1.3.A
August 1993

ESPRIT Basic Research Action 6852 DYANA-2

Dynamic Interpretation of Natural Language

**Universiteit van Amsterdam**
Institute for Logic, Language and Computation (ILLC)

**University of Edinburgh**
Centre for Cognitive Science (ECCS)

**Universität München**
Centrum für Informations- und Sprachforschung (CIS)

**Universitetet i Oslo**
Department of Linguistics and Philosophy (ILF)

**Universität Stuttgart**
Institut für Maschinelle Sprachverarbeitung (IMS)

**Universität Tübingen**
Seminar für Sprachwissenschaft (SfS)

**Rijksuniversiteit Utrecht**
Research Institute for Language and Speech (OTS)

For copies of reports, updates on project activities and other DYANA-related information, contact:

The DYANA-2 Project Administrator
ILLC/Department of Philosophy
University of Amsterdam
Nieuwe Doelenstraat 15
NL-1012 CP Amsterdam

# Contents

## Introduction

The categorial contribution to the 'Grammar architecture' component of DYANA-2 studies systematic extensions of standard Lambek calculus. The objective is to arrive at an empirically articulate and computationally tractable theory of 'linguistic inference'. The theoretical framework is developed within the context of Task 1.1. Task 1.3 confronts the theoretical apparatus with concrete problems of grammar specification, with special emphasis on the dimensions for cross-linguistic variation.

Borrowing terminology from Linear Logic, the type-logical extensions under investigation within this line of research can be classified in three groups.

- Multiplicatives. The standard Lambek connectives embody a purely *concatenative* theory of linguistic composition. But linguistic resources are structured in multiple dimensions — linear order is just one of them. The fine-structure of linguistic composition can be studied in terms of *families* of multiplicative connectives, each with its own specific form of structure-sensitivity of resource management. The families are not considered in isolation, but put together into one encompassing type logic where they freely communicate.
- Additives. Whereas the multiplicative connectives capture the various modes of linguistic composition, the additives provide the means for conjunctive/disjunctive type specifications. From a static perspective, the additive connectives can play a role in structuring the categorial lexicon. From the dynamic perspective, they make it possible to model reasoning with incomplete type information.
- Quantifiers. Extension of the type language with first and second order quantifiers provides a polymorphic type-logical setting to study phenomena of underspecification, unification and generalization — phenomena that form the crucial motivation for 'unification-based' categorial architectures.

The papers in this deliverable relate to the multiplicative and polymorphic generalizations. They represent complementary strategies for extending the basic categorial framework.

The empirical problems at the basis of Morrill's paper are discontinuous dependencies of various sorts. To deal with non-concatenative dependencies, Morrill develops a model in which three algebraic operations live together: plain string concatenation, 'split' string concatenation and a wrapping operation mediating between these two. The type language describing the model consists of residuated families of multiplicative connectives corresponding to the three algebraic operations. Proof-theory for this mixed type system is presented in a labelled deduction format. It is shown how *in situ* binding can be analysed in terms of the interaction of Extraction and Infixation. In situ binding provides an account for quantifier scope ambiguities and for *wh*-binding — with a uniform typing accounting for standard *wh*-extraction and 'pied-piping' constructions. Apart from generalized quantification and *wh*-binding, the paper discusses discontinuous dependencies in idioms, gapping, and object-oriented forms of reflexivisation.

Emms, in contrast, takes a conservative stand with respect to the multiplicatives which, in the case of his contribution, are restricted to the prefixation/suffixation connectives of standard Lambek calculus. To overcome the expressive limitations of a purely concatenative basis, he explores a polymorphic version of the Lambek calculcus, PLCG — the system one obtains by adding second-order type quantifiers to the standard framework. Some logical results about this system are presented in Deliverable R1.1.A. In the present paper PLCG is used to present a polymorphic treatment of ambiguities in coordination and quantification. Conjunctions and generalized quantifiers are assigned schematic types $\forall X.(X\backslash X)/X$ and (systematically ambiguous) $\forall X.X/(np\backslash X)$, $\forall X.(X/np)\backslash X$, respectively. Meaning postulates for the quantifiers and coordination particles are invoked to ensure that the schematic types are interpreted via the familiar generalized quantifier and generalized coordination semantics.

A comparison between the two contributions raises an interesting question with regard to the division of labour between multiplicative and polymorphic extensions of the standard categorial framework. On the one hand, a categorial notion of polymorphic typing is desirable, both within the lexicon, where it offers a means of expressing generalizations in the form of type schemata, and within the context of type inference, where underspecified polymorphic types can dynamically adapt to their environment via instantiation. But on the other hand, it is clear that second-order polymorphism in its full generality is linguistically over-expressive and computationally unmanageable. As Morrill's contribution shows, if one adapts a more discriminating multiplicative theory of linguistic composition, an appeal to polymorphism can be avoided altogether in the case of generalized quantification and *wh*-binding, or restricted to the finite case of conjunctions/disjunctions over basic types.

Further research within the context of Task 2 will have to show whether it is possible to calibrate the interplay between multiplicative and polymorphic extensions in such a way that one can delineate a computationally and linguistically significant fragment of restricted second-order categorial polymorphism, and combine this with a realistic but inherently constrained theory of linguistic composition.

Task 1.3, subtask 3

Polymorphic Treatments

# Some Applications of Categorial Polymorphism

Martin Emms

(Munich University)

# Comments on Emms' "Some Applications of Categorial Polymorphism"

Johan van Benthem
(University of Amsterdam)

# Discontinuity and Pied-Piping in Categorial Grammar

Glyn Morrill

(Universitat Politècnica de Catalunya)