Introduction

Christian Retoré (christian.retore@inria.fr)
Institut National de Recherche en Informatique et Automatique (INRIA)

The papers in this special issue have been selected from a workshop (co-organized with A. Lecomte) entitled *Logical Aspects of Computational Linguistics* that took place in Nancy in September 1995, which was the first of a series of conferences with the same name. Put somewhat pompously, the general theme of the LACL events is nothing less than a reactualization of the connection between logic and grammar, which is as old as these two disciplines. Of course, none of us presumes to define the debate, but rather to see what can be achieved with such a logical view of grammar, in particular from a computational perspective. Therefore this research program — see [1] for more details and references — includes both linguistic modeling and mathematical developments.

Authors have been encouraged to submit the most recent developments of their research, rather than to stick to their well-prepared talks. This resulted in the present issue, which, due to the hazards of submission and selection, does not cover all the themes present in the *LACL* conference series. This is the reason why we call this issue *Recent advances in logical and algebraic approaches to grammar.* Papers focus on the relations between grammar, logic and algebra, with the Lambek calculus playing a central rôle. The linguistic formalisms involved are Categorial Grammars and Definite Clause Grammars; logic (intuitionistic, linear, modal) is mainly considered from a proof-theoretical perspective but some truth-value models are also considered; the kind of algebra involved is group theory, semi-group theory, lattices over monoids. This may seem a wide spectrum, but either explicitly or implicitly, the logical formalization of the kind of categorial grammar known as the Lambek calculus is omnipresent, since each paper may be viewed as a connection between the Lambek calculus and another area of logic or linguistics.

One of the main advantage of the logical description of categorial grammar provided by Lambek is its neat interface with semantics, especially with the kind of semantics that Montague formalized (see e.g. [9] for a survey). Indeed the connection between a Lambek-style syntactic analysis and a lambda-term expressing the predicate-calculus formula is as close as possible: the type homomorphism from syntactic to semantic types is totally uniform, and, via the Curry-Howard isomorphism, each derivational step of the categorial grammar (an inference rule of the

Lambek calculus) is mapped onto a lambda term construction step. Thus, out of the lambda terms representing the semantics of the lexical items, one sees the meaning of phrases arising step by step according to the syntactic rules. This tight connection to semantics was very convincing, and historically unsurprising [3], but, because of the generative and descriptive limits of stricto sensu Lambek calculus, there was a need to relate this logical island to more developed approaches to grammar on the one hand and to other logics on the other hand.

For instance the revival of categorial grammars in the eighties is mainly due to advances in logic. Researchers such as Moortgat [8] and van Benthem [16] have explored the relationship between Lambek calculus and modal logic, thus extending its abilities while remaining inside the same parsing-as-deduction paradigm — which is the source of the pleasant interface with semantics. The basic idea is that structural modalities should be able to explicitly control the resource sensitivity of the logic: one does not want to always have associativity, commutativity, etc. but rather a controlled use of such equivalences. *Categorial inference and modal logic* by Natasha Kurtonina shows that the Lambek calculus can be faithfully embedded into unimodal temporal logic. In other words, the resource sensitivity of the Lambek calculus can be described in a different kind of logic, the advantages of which may benefit to the Lambek calculus, like frame semantics and Kripke models.

As said above the Curry-Howard isomorphism plays a central rôle in the connection between Montagovian semantics and categorial grammars. According to this isomorphism, lambda-terms may be viewed as proofs in intuitionistic logic; it is therefore quite natural to describe categorial grammars in the formalism of intuitionistic type theory [7] where intuitionistic logic is handled in the most natural way: intuitionistic type theory can provide a common framework for dealing both with syntax and semantics. *Syntactic calculus and dependent types* by Aarne Ranta encodes the syntactic calculus in constructive type theory. He shows how dependent types are needed in particular to handle interpretation and syntactic structures simultaneously. Indeed, simply typed lambda calculus can only represent proofs in intuitionistic propositional logic ; predicate logic needs a dependently typed calculus, because e.g. universal quantification corresponds to the type $(x : A)B(x)$ and not to $A \to B$, which corresponds to implication.

In the eighties Girard invented linear logic [4, 5, 14], with its deep proof-theoretical roots. Linear logic has been of great importance for understanding how logic can handle resource management, and the fine grained functioning of intuitionistic and classical logic. Regarding categorial grammars, the Lambek calculus appears to be exactly intuitionistic non-commutative linear logic (see [12] for a survey). Therefore linear logic offers new ways of extending Lambek calculus while staying within a proof theoretical framework. One may also expect that the linear logic view of the Lambek calculus will make clearer its relation to classical or intuitionistic logic, which are the standard logics for semantics. Furthermore linear logic allows a radically new syntax for proofs (that is syntactic analysis): proof nets. Roughly speaking these graphs are to linear logic what natural deduction trees are to intuitionistic logic. They consist in a family of trees (expressing the logical structure of the conclusions) plus some axioms linking dual leaves; in order to represent proofs, these graphs must satisfy a global correctness criterion. From a linguistic viewpoint, an axiom between a positive atom a of a compound category A and a negative atom $-a$ of a compound category B expresses that the demand of a by A is fulfilled by the $-a$ of B, i.e. corresponds to the consumption of a valency. As opposed to sequent calculus, this syntax eliminates irrelevant differences between proofs, like the order of application of independent rules. So from a proof theoretical viewpoint they are exciting structures, but they should also be appealing from a linguistic viewpoint: if we consider a grammatical analysis to be a proof then getting closer to the essence of a proof should bring some linguistic insight. *Proof nets and the complexity of processing center-embedded constructions* by Mark Johnson illustrates how interesting proof nets are from a linguistic viewpoint. He addresses a performance question within Lambek calculus: using proof nets as parse structures, he is able to measure the complexity of processing center-embedded relative clauses. This connection between psycholinguistics and categorial grammars — rare as far as I know — seems highly promising. *A new correctness criterion for multiplicative cyclic linear logic* by Michele Abrusci and Elena Maringelli provides a simple and elegant description of proof nets for cyclic linear logic. This logical calculus is the simplest classical (with an involutive negation) extension of the Lambek calculus, and their result also applies to the Lambek calculus by asking formulas to be intuitionistic or polarised.

But, when one speaks of logic and proofs, one asks for models. Given the grammatical motivations we have in mind, models naturally take place within *monoids*: a syntactic category is interpreted by the set of strings of this category, and this is a standard model for the Lambek calculus, as exemplified in this issue in the paper by Kurtonina (see [2] for a survey). The part of algebra known as *formal language theory* [13] is precisely devoted to the study of monoids, and is certainly the best studied part of mathematical linguistics. In the eighties phrase-structure grammars were generalized to unification grammars — thus meeting logic in quite a different way [10]. However the relation between phrase-structure grammars and categorial grammars still remains problematic, an important exception being the result of Mati Pentus in 1992 who established the weak equivalence between context-free grammars and Lambek grammars; unsurprisingly he makes use of a string model for the Lambek calculus, but in a group rather than in a monoid (see his recent article [11]). *Group theory and computational linguistics* by Marc Dymetman is an innovative way of linking categorial grammars and unification grammars. Indeed the author, placing himself into what may be viewed as a model from a categorial viewpoint and by making use of standard constructions of group theory such as conjugacy and normality, reduces grammatical derivations to proving equalities in a group. Although the exact relationship with models of the Lambek calculus has to be made precise, the potential connections are intriguing. And of course the purely algebraic treatment of unification grammar is by itself a significant result.

Let me end up with some general references and many thanks to the authors and referees.

References

1. Patrick Blackburn, Marc Dymetman, Alain Lecomte, Aarne Ranta, Christian Retoré and Eric Villemonte de la Clergerie Logical Aspects of Computational Linguistics: an introduction. In C. Retoré, editor. *Logical Aspects of Computational Linguistics, LACL'96*, volume 1328 of *LNCS/LNAI*. Springer-Verlag, 1997.
2. Wojciech Buszkowski. Mathematical linguistics and proof theory. In van Benthem and ter Meulen [15], chapter 12, pages 683–736.
3. Claudia Casadio. Semantic categories and the development of categorial grammars. In R. Oehrle, E. Bach, and D. Wheeler, editors. *Categorial Grammars and Natural Language Structures*, Reidel, Dordrecht, 1988, pages 95–124.
4. Jean-Yves Girard. Linear logic. *Theoretical Computer Science*, 50(1):1–102, 1987.
5. Jean-Yves Girard. Linear logic: its syntax and semantics. In J.-Y. Girard, Y. Lafont, and L. Regnier, editors. *Advances in Linear Logic*, volume 222 of *London Mathematical Society Lecture Notes*, Cambridge University Press, 1995, pages 1–42.
6. Joachim Lambek. The mathematics of sentence structure. *American mathematical monthly*, pages 154–170, 1958.
7. Per Martin-Löf. *Intuitionistic Type Theory*. Bibliopolis, Napoli, 1984. (Lecture Notes by G. Sambin).
8. Michael Moortgat. Categorial type logic. In van Benthem and ter Meulen [15], chapter 2, pages 93–177.
9. Barabara Partee with Herman Hendriks. Montague grammar. In van Benthem and ter Meulen [15], chapter 1, pages 5–92.

10. Fernando C.N. Pereira and Stuart M. Shieber. *Prolog and Natural-Language Analysis*, volume 10 of *CSLI Lecture Notes*. CSLI, 1987. (distributed by Cambridge University Press).
11. Mati Pentus. Product-free lambek calculus and context-free grammars. *Journal of Symbolic Logic*, 62(2):648–660, 1997.
12. Christian Retoré. Calcul de lambek et logique linéaire. *Traitement Automatique des Langues*, 37(2):39–70, 1996.
13. G. Rozenberg and A. Salomaa, editors. *Handbook of Formal Languages*. Springer Verlag, Berlin, 1997.
14. Anne Sjerp Troelstra. *Lectures on Linear Logic*, volume 29 of *CSLI Lecture Notes*. CSLI, 1992. (distributed by Cambridge University Press)
15. J. van Benthem and A. ter Meulen, editors. *Handbook of Logic and Language*. North-Holland Elsevier, Amsterdam, 1997.
16. J. van Benthem. *Language in Action: Categories, Lambdas and Dynamic Logic*. North-Holland Elsevier, Amsterdam, 1991.