IN PARTNERSHIP WITH:
CNRS

Université de Lorraine

Activity Report 2014

Project-Team SEMAGRAMME

Semantic Analysis of Natural Language

IN COLLABORATION WITH: Laboratoire lorrain de recherche en informatique et ses applications (LORIA)

RESEARCH CENTER
Nancy - Grand Est

THEME
Language, Speech and Audio

Table of contents

Project-Team SEMAGRAMME

Keywords: Logics, Natural Language, Semantics, Linguistics

Creation of the Team: 2011 January 01*, updated into Project-Team:* 2013 July 01.

1. Members

Research Scientists
Philippe de Groote [Team leader, Inria, Senior Researcher]
Bruno Guillaume [Inria, Researcher]
Sylvain Pogodalla [Inria, Researcher]

Faculty Members
Maxime Amblard [Univ. Lorraine, Associate Professor]
Guy Perrier [Univ. Lorraine, Professor, HdR]
Karën Fort [Univ. Lorraine, ATER until august 2014]

PhD Students
Jirí Marsík [Inria, CORDI-S]
Aleksandre Maskharashvili [Inria, granted by ANR POLYMNIE project]
Sai Qian [Univ. Lorraine, until Nov 2014]

Post-Doctoral Fellow
Can Baskent [Inria]

Administrative Assistants
Antoinette Courrier [CNRS]
Sylvie Musilli [Univ. Lorraine]
Céline Simon [Inria]

Others
Nicolas Anglade [Inria, intern INPL, from Jun 2014 until Sep 2014]
Nicolas Barrandon [Inria, intern Univ. Paris Sud, from Jun 2014 until Jul 2014]
Sarah Blind [Inria, intern Univ. Lorraine, until Jun 2014]
Mikael Monet [Inria, intern INPL, from Jun 2014 until Sep 2014]
Pascal Pierrot [Univ. Lorraine, intern INPL, from Jun 2014 until Sep 2014]
Valentin Stern [Inria, intern Univ. Lorraine, from Apr 2014 until Jun 2014]

2. Overall Objectives

2.1. Overall objectives

Computational linguistics is a discipline at the intersection of computer science and linguistics. On the theoretical side, it aims to provide computational models of the human language faculty. On the applied side, it is concerned with natural language processing and its practical applications.

From a structural point of view, linguistics is traditionally organized into the following sub-fields:

- Phonology, the study of language abstract sound systems.
- Morphology, the study of word structure.
- Syntax, the study of language structure, i.e., the way words combine into grammatical phrases and sentences.
- Semantics, the study of meaning at the levels of words, phrases, and sentences.
- Pragmatics, the study of the ways in which the meaning of an utterance is affected by its context.

Computational linguistics is concerned by all these fields. Consequently, various computational models, whose application domains range from phonology to pragmatics, have been developed. Among these, logic-based models play an important part, especially at the "higher" levels.

At the level of syntax, generative grammars [48] may be seen as basic inference systems, while categorial grammars [65] are based on substructural logics specified by Gentzen sequent calculi. Finally, model-theoretic grammars [77] amount to sets of logical constraints to be satisfied.

At the level of semantics, the most common approaches derive from Montague grammars, [67], [68], [69] which are based on the simply typed λ-calculus and Church's simple theory of types [49]. In addition, various logics (modal, hybrid, intensional, higher- order...) are used to express logical semantic representations.

At the level of pragmatics, the situation is less clear. The word *pragmatics* has been introduced by Morris [71] to designate the branch of philosophy of language that studies, besides linguistic signs, their relation to their users and the possible contexts of use. The definition of pragmatics was not quite precise, and for a long time several authors have considered (and some authors are still considering) pragmatics as the wastebasket of syntax and semantics [40]. Nevertheless, as far as discourse processing is concerned (which includes pragmatic problems such as pronominal anaphora resolution), logic-based approaches have also been successful. In particular, Kamp's Discourse Representation Theory [63] gave rise to sophisticated 'dynamic' logics [58]. The situation, however, is less satisfactory than it is at the semantic level. On the one hand, we are facing a kind of logical "tower of Babel". The various pragmatic logic-based models that have been developed, while sharing underlying mathematical concepts, differ in several respects and are too often based on *ad hoc* features. As a consequence, they are difficult to compare and appear more as competitors than as collaborative theories that could be integrated. On the other hand, several phenomena related to discourse dynamics (e.g., context updating, presupposition projection and accommodation, contextual reference resolution...) are still lacking deep logical explanations. We strongly believe, however, that this situation can be improved by applying to pragmatics the same approach Montague applied to semantics, using the standard tools of mathematical logic.

Accordingly:

> *The overall objective of the Sémagramme project is to design and develop new unifying logic-based models, methods, and tools for the semantic analysis of natural language utterances and discourses. This includes the logical modelling of pragmatic phenomena related to discourse dynamics. Typically, these models and methods will be based on standard logical concepts (stemming from formal language theory, mathematical logic, and type theory), which should make them easy to integrate.*

The project is organized along three research directions (i.e., *Syntax-semantics interface*, *Discourse dynamics*, and *Common basic resources*), which interact as explained in the following paragraphs.

2.1.1. Syntax-semantics interface

The Sémagramme project intends to focus on the semantics of natural languages (in a wider sense than usual, including some pragmatics). Nevertheless, the semantic construction process is syntactically guided, that is, the constructions of logical representations of meaning is based on the analysis of the syntactic structures. We do not want, however, to commit ourselves to such or such specific theory of syntax. Consequently, our approach should be based on an abstract generic model of the syntax-semantic interface.

Here, an important idea of Montague comes into play, namely, the "homomorphism requirement": semantics must appear as a homomorphic image of syntax. While this idea is almost a truism in the context of mathematical logic, it remains challenged in the context of natural languages. Nevertheless, Montague's idea has been quite fruitful, especially in the field of categorial grammars, where van Benthem showed how syntax ans semantics could be connected using the Curry-Howard isomorphism [86]. This correspondence is the keystone of the syntax-semantics interface of modern type-logical grammars [70]. It also motivated the definition of our own Abstract Categorial Grammars [81].

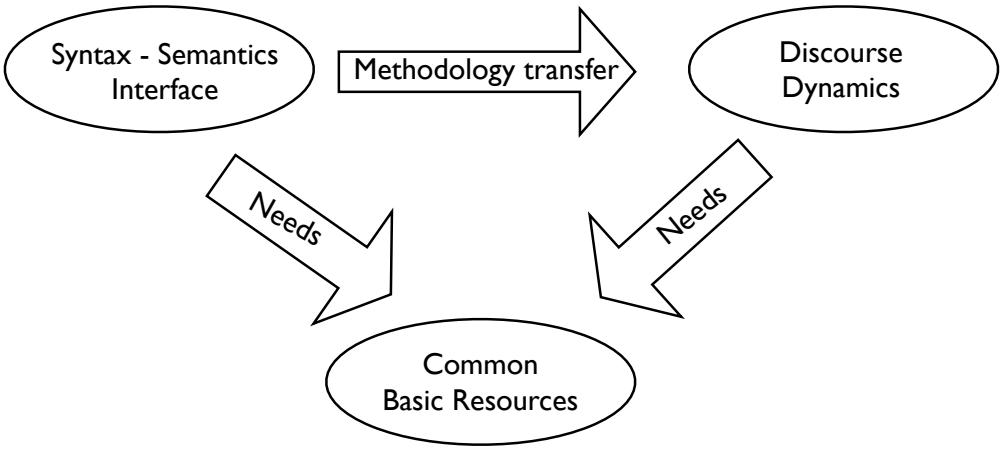

Figure 1.

Technically, an Abstract Categorial Grammar simply consists of a (linear) homomorphism between two higher-order signatures. Extensive studies have shown that this simple model allows several grammatical formalisms to be expressed, providing them with a syntax-semantics interface for free. [82], [84], [85], [73], [64], [76].

We intend to carry on with the development of the Abstract Categorial Grammar framework. At the foundational level, we will define and study possible type theoretic extensions of the formalism, in order to increase its expressive power and its flexibility. At the implementation level, we will continue the development of an Abstract Categorial Grammar support system.

As said above, to consider the syntax-semantics interface as the starting point of our investigations allows us not to be committed to some specific syntactic theory. The Montagovian syntax-semantics interface, however, cannot be considered to be universal. In particular, it does not seem to be that well adapted to dependency and model-theoretic grammars. Consequently, in order to be as generic as possible, we intend to explore alternative models of the syntax-semantics interface. In particular, we will explore relational models where several distinct semantic representations can correspond to a same syntactic structure.

2.1.2. Discourse dynamics

It is well known that the interpretation of a discourse is a dynamic process. Take a sentence occurring in a discourse. On the one hand, it must be interpreted according to its context. On the other hand, its interpretation affects this context, and must therefore result in an updating of the current context. For this reason, discourse interpretation is traditionally considered to belong to pragmatics. The cut between pragmatics and semantics, however, is not that clear.

As we mentioned above, we intend to apply to some aspects of pragmatics (mainly, discourse dynamics) the same methodological tools Montague applied to semantics. The challenge here is to obtain a completely compositional theory of discourse interpretation, by respecting Montague's homomorphism requirement. We think that this is possible by using techniques coming from programming language theory, in particular, continuation semantics [79], [41], [42], [78] and the related theories of functional control operators [54], [55].

We have indeed successfully applied such techniques in order to model the way quantifiers in natural languages may dynamically extend their scope [83]. We intend to tackle, in a similar way, other dynamic phenomena (typically, anaphora and referential expressions, presupposition, modal subordination...).

What characterize these different dynamic phenomena is that their interpretations need information to be retrieved from a current context. This raises the question of the modeling of the context itself. At a foundational level, we have to answer questions such as the following. What is the nature of the information to be stored in the context? What are the processes that allow implicit information to be inferred from the context? What are the primitives that allow a context to be updated? How does the structure of the discourse and the discourse relations affect the structure of the context? These questions also raise implementation issues. What are the appropriate datatypes? How can we keep the complexity of the inference algorithms sufficiently low?

2.1.3. Common basic resources

Even if our research primarily focuses on semantics and pragmatics, we nevertheless need syntax. More precisely, we need syntactic trees to start with. We consequently need grammars, lexicons and parsing algorithms to produce such trees. During the last years, we have developped the notion of interaction grammar [59] as a model of natural language syntax. This includes the development of grammar for French, [72] together with morpho-syntactic lexicons. We intend to continue this line of research and development. In particular, we want to increase the coverage of our French grammar, and provide our parser with more robust algorithms.

Further primary resources are needed in order to put at work a computational semantic analysis of utterances and discourses. As we want our approach to be as compositional as possible, we must develop lexicons annotated with semantic information. This opens the quite wide research area of lexical semantics.

Finally, when dealing with logical representations of utterance interpretations, the need for inference facilities is ubiquitous. Inference is needed in the course of the interpretation process, but also to exploit the result of the interpretation. Indeed, an advantage of using formal logic for semantic representations is the possibility of using logical inference to derive new information. From a computational point of view, however, logical inference may be highly complex. Consequently, we need to investigate which logical fragments can be used efficiently for natural language oriented inference.

3. Research Program

3.1. Foundations

The Sémagramme project relies on deep mathematical foundations. We intend to develop models based on well-established mathematics. We seek two main advantages from this approach. On the one hand, by relying on mature theories, we have at our disposal sets of mathematical tools that we can use to study our models. On the other hand, developing various models on a common mathematical background will make them easier to integrate, and will ease the search for unifying principles.

The main mathematical domains on which we rely are formal language theory, symbolic logic, and type theory.

3.1.1. Formal language theory

Formal language theory studies the purely syntactic and combinatorial aspects of languages, seen as sets of strings (or possibly trees or graphs). Formal language theory has been especially fruitful for the development of parsing algorithms for context-free languages. We use it, in a similar way, to develop parsing algorithms for formalisms that go beyond context-freeness. Language theory also appears to be very useful in formally studying the expressive power and the complexity of the models we develop.

3.1.2. Symbolic logic

Symbolic logic (and, more particularly, proof-theory) is concerned with the study of the expressive and deductive power of formal systems. In a rule-based approach to computational linguistics, the use of symbolic logic is ubiquitous. As we previously said, at the level of syntax, several kinds of grammars (generative, categorial...) may be seen as basic deductive systems. At the level of semantics, the meaning of an utterance is capture by computing (intermediate) semantic representations that are expressed as logical forms. Finally, using symbolic logics allows one to formalize notions of inference and entailment that are needed at the level of pragmatics.

3.1.3. Type theory and typed λ-calculus

Among the various possible logics that may be used, Church's simply typed λ-calculus and simple theory of types (a.k.a. higher-order logic) play a central part. On the one hand, Montague semantics is based on the simply typed λ-calculus, and so is our syntax-semantics interface model. On the other hand, as shown by Gallin, [56] the target logic used by Montague for expressing meanings (i.e., his intensional logic) is essentially a variant of higher-order logic featuring three atomic types (the third atomic type standing for the set of possible worlds).

4. Application Domains

4.1. Introduction

Our applicative domains concern natural language processing applications that rely on a deep semantic analysis. For instance, one may cite the following ones:

- textual entailment and inference,
- dialogue systems,
- semantic-oriented query systems,
- content analysis of unstructured documents,
- text transformation and automatic summarization,
- (semi) automatic knowledge acquisition.

However, if the need for semantics seems to be ubiquitous, there is a challenge in finding applications for which a deep semantic analysis results in a real improvement over non semantic-based techniques.

4.2. Text Transformation

Text transformation is an application domain featuring two important sub-fields of computational linguistics:

- parsing, from surface form to abstract representation,
- generation, from abstract representation to surface form.

Text simplification or automatic summarization belong to that domain.

We aim at using the framework of Abstract Categorial Grammars we develop to this end. It is indeed a reversible framework that allows both parsing and generation. Its underlying mathematical structure of λ-calculus makes it fit with our type-theoretic approach to discourse dynamics modeling. The ANR project Polymnie (see section 7.2.1.1) is especially dedicated to this aim.

5. New Software and Platforms

5.1. ACG Development Toolkit

Participants: Sylvain Pogodalla [correspondent], Philippe de Groote, Jirí Marsík.

In order to support the theoretical work on ACG, we have been developing a support system. The objectives of such a system are twofold:

1. To make possible to implement and experiment grammars the modeling of linguistic phenomena.
2. To make possible to implement and experiment results related to the ACG formalisms. Such results can concern parsing algorithms, type extensions, language extensions, etc.

The ACG Development toolkit development effort is part of the POLYMNIE project (see Section 7.2.1.1). It will support the experimentation and evaluation parts of the project.

The current version of the ACG development toolkit prototype [1] is 1.1. It focuses on providing facilities to develop grammars. To this end, the type system currently implemented is the linear core system plus the (non-linear) intuitionistic implication, and a special attention has been paid to type error management. Since 1.0b released in Feb. 2014, ACGtk allows for transformations both from abstract terms to object terms, and from object terms to abstract terms (ACG parsing). The parsing algorithm follows [64]'s method which is being implemented for second-order ACGs. It is based on a translation of ACG grammars into Datalog programs and is well-suited to fine-grained optimization.

However, since we are interested not only by recognizability (hence whether some fact is provable) but also by the parsing structure (hence the proof), the Datalog solver has been adapted to produce not only yes/no answer to queries, but also all the proofs of the answers to the queries. The next steps concern optimization and efficiency. Note however that in the general case, the decidability of translating an object term to an abstract one is still an open problem.

We also have enriched the ACG development toolkit with graphical output. The new module includes a small functional OCaml library for manipulating images which enables users to customize the rendering of formulas as pictures.

The ACGtk has been made available as an OPAM (OCaml Package Manager) package. [2]

5.2. Grew

Participants: Bruno Guillaume [correspondent], Guy Perrier.

Grew (http://grew.loria.fr) is a Graph Rewriting tool dedicated to applications in NLP. It is freely-available and it is developed using the InriaGforge platform (http://gforge.inria.fr/projects/semagramme/).

Grew takes into account confluent and non-confluent graph rewriting and it includes several mechanisms that help to use graph rewriting in the context of NLP applications (built-in notion of feature structures, parametrization of rules with lexical information).

In 2014, an online version (http://talc2.loria.fr/grew/) of the tool based on the matching part was developped to illustrate its use (it is not possible to modify graphs). The user gives a pattern (eventually with some negative constraints) and Grew searches in a corpus the occurences on the given pattern in: the French corpus Sequoia is available (two versions are available: one containing surface annotation and one with deep annotation 6.4) and the German corpus Tiger is also avalaible for online pattern search.

5.3. Leopar

Participants: Bruno Guillaume [correspondent], Guy Perrier.

Leopar is a parser for natural languages which is based on the formalism of Interaction Grammars [59]. It is open-source (under the CECILL License http://www.cecill.info) and it is developed using the InriaGforge platform (http://gforge.inria.fr/projects/semagramme/).

[1] Available at http://acg.gforge.inria.fr with a CeCILL license.
[2] https://opam.ocaml.org/packages/acgtk/acgtk.1.1/

The main features of current version of the software are:

- automatic parsing of a sentence or a set of sentences,
- dependency and parse-tree representation of sentences,
- interactive parsing (the user chooses the couple of nodes to merge),
- visualization of grammars produced by XMG-2 or of sets of description trees associated to some word in the linguistic resources.

In 2014, a new conversion from parse-tree representation to dependency representation was implemented to take benefit of the linguistic principles that were defined and used in [36].

5.4. ZombiLingo

Participants: Bruno Guillaume [correspondent], Karën Fort.

Zombilingo (http://zombilingo.loria.fr) is a prototype of a GWAP where gamers have to give linguistic information about the syntax of French natural language sentence (see 6.6 for more details).

5.5. Other developments

Participants: Maxime Amblard [correspondent], Bruno Guillaume.

Main topics: data managment, disfluencies and dependency

- Dep2pict (http://dep2pict.loria.fr) is a program for drawing graphical representation of dependency structures of natural language sentences. An online version is available (http://wikilligramme.loria.fr/doku.php/dep2pict:demo). In 2014, the Dep2pict was modified to take into account the modified format mixing surface and deep syntactic information used in deep-sequoia 6.4.
- A management chain of the transcriptions of interviews for the SLAM project which productes of a full anonymized randomized version of the resources.
- A program based on Distagger (disfluences) and MElt (POS and lemma) and proposes different repartition analyses.

6. New Results

6.1. Generation

G-TAG [52], [66] is a Tree Adjoining Grammar (TAG) based formalism which was specifically designed for the task of text generation. Contrary to TAG, the derivation structure becomes primary, as pivot between the conceptual representation and the surface form. This is a shared feature with the encoding of TAG into ACG. Laurence Danlos (Alpage Inria project), Aleksandre Maskharashvili, and Sylvain Pogodalla have shown how to recast the G-TAG formalism into ACG, relying on the reversibility properties of the later [17], [16], [18].

6.2. Discourse Grammar

Laurence Danlos (Alpage Inria project), Aleksandre Maskharashvili, and Sylvain Pogodalla have presented a method to interface a sentential grammar and a discourse grammar. It offers both a smooth integration of the two grammars without using an intermediate processing step, and the possibility to build discourse structures that are direct acyclic graphs (DAG) and not only trees. The analysis is based on a Tree-Adjoining Grammar (TAG) approach to discourse: Discourse Synchronous TAG (D-STAG) [50], [51], and uses an encoding of TAG into ACG. This allows for expressing a higher-order semantic interpretation that enables building DAG discourse structures, and for smoothly integrating the sentential and the discourse grammar thanks to the modular capability of ACG. All the examples may be run and tested with the the ACGtk (submitted).

6.3. Large Scale Grammatical Resources

Guy Perrier wrote a complete documentation [36] on FRIGRAM [3] a French grammar with a large coverage, written in the formalism of Interaction Grammars [59]. The different chapters of the 257 pages of documentation correspond to the different parts of speech in French. At the end, two chapters are dedicated to two specific phenomena: extraction (relative, interrogative and cleft clauses) and coordination, which is presented in common with punctuation because of their proximity.

6.4. Deep Syntax Annotation of the Sequoia French Treebank

Marie Candito, Guy Perrier, Bruno Guillaume, Corentin Ribeyre, Karën Fort, Djamé Seddah and Eric de la Clergerie annotated the Sequoia French Treebank with deep syntax dependencies [14].

The Sequoia French Treebank [47] is a 3.100 sentences treebank covering several domains (news, medical, europarl and fr-wikipedia). It is freely available and has already been annotated with surface dependency representations.

The participants in the project have defined a deep syntactic representation scheme for French, built from the surface annotation scheme of the Sequoia corpus and abstracting away from it [28]. This scheme expresses the grammatical relations between content words. When these grammatical relations take part into verbal diatheses, the diatheses are considered as resulting from redistributions from the canonical diathesis, which is retained in the annotation scheme.

The goal is to obtain a freely available corpus, which will be useful for corpus linguistics studies and for training deep analyzers to prepare semantic analysis.

The different steps of the annotation process were conducted in a collaborative way. As the members of the project are located in two different French towns (Paris and Nancy), they decided to produce a complete annotation of the TreeBank in both towns and to collaboratively adjudicate the two results.

Each team separately produced an initial annotated version of the mini reference. The final version, resulting from several iterations and adjudications, is available [4].

6.5. Exploitation of the LVF (Lexicon of French Verbs)

Bruno Guillaume, Karën Fort, Guy Perrier and Paul Bédaride have worked on the LVF [53] ("Lexique des Verbes du Français", Lexicon of French Verbs). This large lexicon was build by two French linguists, Jean Dubois and Françoise Dubois-Charlier and contains detailed linguistic information about 12.308 lemmas of French verbs. The work presented in [21] describes experiments aiming at mapping the LVF to DICOVALENCE [87]. The two resources (LVF and DICOVALENCE) were built by linguists, based on very different theories, which makes a direct mapping nearly impossible. In the current work, we focus on the linguistic examples given in LVF. These examples are not sentences that can be parsed directly; the first part of the work was to express examples as real natural language sentence. It is then possible to use FRILEX, a Natural Language Processing lexicon based on DICOVALENCE to parse corrected examples given with LVF entries. This results in an automatic partial mapping of LVF entries against DICOVALENCE entries.

6.6. Game With A Purpose

Crowdsourcing is nowadays a way of constructing linguistic resources which is more and more used. In the crowdsourcing area, one of the way to motivate a large amount of people to contribute to a project is to present it as a game. Games used in this particular way are called GWAPs (Game With A Purpose).

In Natural Language Processing, examples of GWAP are "Phrase detective" where games are asked to resolve anaphora in English texts and "JeuDeMots" where gamers have to given lexical terms related to a term given by the system (the goal is to build a semantic networks of French lexical items).

[3] http://wikilligramme.loria.fr/doku.php/frig:frig
[4] https://deep-sequoia.inria.fr

Karën Fort and Bruno Guillaume worked on the definition of a GWAP to help construction of syntactically annotated corpora. With a student (Hadrien Chastant), they presented in April, the design of ZombiLingo [20], a GWAP that allows for the dependency syntax annotation of French corpora. The main aspects of this work are to explain: how to deal with the complexity of the task, how to motivate gamers to contribute and how to ensure that a large numbers of gamers will help to produce an high quality linguistic resource.

With another student (Valentin Stern), a first prototype was built. This first version implements only a part of the mechanisms described in the previous work and it is used as a proof-of-concept of a future game. This prototype was presentend at the TALN conference in July [27].

6.7. Supertagging

Guillaume Bonfante, Bruno Guillaume, Mathieu Porey and Guy Perrier wrote a book chapter [30] "Supertagging with constraints". This chapter makes a survey of the results obtained in previous publications about supertagging based on polarities [46] and based on the companionship principle [45]. The last section of the chapter presents a new application of the companionship principle to the TAG formalism and presents some experimental results.

6.8. Modelling Semantic Phenomena

Despite the valuable insights yielded by the classical theories of discourse semantics, there is a wide range of exceptional phenomena that they fail to address, e.g., anaphora under double negation and modality. Concentrating on these two exceptions, Sai Qian, Philippe de Groote and Maxime Amblard provide a corresponding adaptation of TTDL for each case. Briefly speaking, for the problem of double negation, they propose to encapsulate both the affirmative representation and the negative representation of an expression in its semantics. Negation is treated as an operation which switches the positions of the two representations. Thus a second negation will switch the positions again as if no negation had ever occurred. In this way, a double negation can be eliminated and the desired referent accessibility is modelled. As for anaphora under modality, they propose to enrich the TTDL left context with the notion of modal base, which is proposed by Kratzer. The possible world model is integrated in the semantic representation as well. Moreover, they show how the different adaptations could work in an unified framework, [75].

6.9. Quantification in event semantics

Yoad Winter (Utrecht University) has given a type-logical account of quantification in event semantics.

It has been observed in the literature that Davidson's event semantics does not combine smoothly with Montague's compositional semantics. The difficulty comes from a possibly bad interaction between event existential closure, on the one hand, and quantification, negation, or conjunction, on the other hand. In a recent publication, Winter and Zwarts provide a solution to this problem. Winter and de Groote elaborate on this solution. In particular, they provide a treatment of quantified adverbial modifiers, which was absent from Winter and Zwarts, [19].

6.10. Pragmasemantic with Effects and Handlers

Jiří Maršík and Maxime Amblard have explored the feasibility of theories of side effects of programming languages in the study of natural language semantics and pragmatics [23]. In the approach that we are developing, the denotations we assign to fragments natural language are effectful computations. To demonstrate on an example, if we was to treat dynamics, then instead of changing the type of sentence denotations from o to c -> o * c, where c is the type of discourse contexts and o is the type of propositions, we would treat sentence denotations as effectful computations of type o that study and modify the context using effectful operations. This explicit distinction between 'result' and 'effects' brings to mind Stalnaker's distinction between 'content' and 'context'.

The motivation for this approach is to make it easier to compose multiple pragmasemantic phenomena by being allowed to put their effects aside. So far, a small prototype handling dynamics, presuppositions and some of their interactions is under development.

6.11. Mining Texts at discourse level

Linguistic discourse refers to the meaning of large chunks of text, from phrases to whole documents. It could be very useful for guiding attempts at text mining, which focus on document selection, document summarization, or other knowledge extraction goals. Hence the aim of this work is to apply Knowledge Discovery in Databases (KDD) methods to texts annotated with discourse information. Maxime Amblard with Yannick Toussaint (Orpailleur team) and Sara van de Moosdijk (master 2 intern) approach the problem by extracting discourse relations using unsupervised methods, which are then used to construct a knowledge model with Formal Concept Analysis (FCA). Pattern Structures (PS), an advancement in FCA, allow for the modelling of complex data. Our method is applied to a corpus of medical articles compiled from PubMed. This medical data is enhanced with concepts from the UMLS MetaThesaurus combined with the UMLS Semantic Network to serve as an ontology for Pattern Structure classification. The results show that despite having a large amount of noise, the method is promising and could be applied to other domains than the medical domain. We explore the pitfalls and suggest ways in which the process could be improved (Submission under review).

6.12. Exploring real datas

Maxime Amblard explored the use of formal framework for modelling transcription of real interviews, in particular one involves in the SLAM project with schizophrenics. Schizophrenia is well-known among mental illnesses for the severity of the thought disorders it involves, and for their widespread and spectacular manifestations: from deviant social behavior to delusion, not to mention affective and sensitive distortions. The goal of our interdisciplinary work is to (i) analyze linguistic troubles in conversational contexts in which one of the speakers is schizophrenic, (ii) construe how the concept of rationality and logicality may apply to them, and (iii) propose a formal representation about this specific manifestation. Maxime Amblard, Sylvain Pogodalla and Karen Fort propose surveys on past results [35], [29].

Maxime Amblard and Karen Fort have studied experiments they led concerning disfluencies in the discourse of schizophrenic patients (in remediation). These experiments are part of a larger study dealing with other levels of linguistic analysis, that could eventually help identifying clues leading to the diagnostic of the disease. This study largely relies on natural language processing tools, which allow for the rapid processing of massive textual data (here, more than 375,000 words). The first phase of the study, which they present confirmed the correlation between schizophrenia and the number of disfluences appearing in the discourse [25]. Moreover they have discussed ethical issues on the corpus with others [26].

6.13. Paraconsistency and Inconsistency-Friendly Logics

Paraconsistent logic is a family of formal systems in which the law of contradiction fails. In such systems, from an inconsistent set, *not* everything follows.

Can Baskent has studied such logical systems and their connections to formal linguistics within the framework of game theory. First, he observed how a game theoretical semantics can be given for some paraconsistent logics [43] . The advantage of game semantics is that it simply reflects the parsing tree of logic, and furthermore presents a semantical structure that uses elements from game theory. Such a study also requires an in-depth study of various paraconsistent logics, and their semantical structures [13]. Such a study requires some understanding of point-set topology, and its relation to logic.

Moreover, paraconsistent logics relate to dynamic logics as well. The logical model defines characterises how dynamic epistemic modalities, which are familiar from multi-agent systems, work [13]. This helps us understand how multi-agent interactions in an inconsistent model work in a sound way.

Another interesting way of seeing how inconsistency-friendly logics work is to consider them within the framework of game theory [37]. Game theory, similar to multi-agent systems, studies the intelligent and rational interaction of decision makers/agents. Yet, it suffers from various paradoxes. Such paradoxes are important from a computational semantical point of view. If paraconsistency is the most suitable tool to analyse paradoxes, then game theoretical paradoxes are not exceptions [37].

The technical work always needs to be supplemented by some conceptual work. Granted, paraconsistent logics find their ways in various philosophical and semantical issues, yet their computational analysis usually falls short. In [44], we discussed the connection between paraconsistent logics and Hintikka's interrogative models. These models have been developed by Hintikka, a pioneer of epistemic logic, and have been properly analysed from paraconsistent perspectives. If inquiry and questioning needs to be accounted for computationally, a paraconsistent approach will be an appropriate tool as well. Similarly, [39] discusses paraconsistency and its connection to social software. *Social Software* is a field conceived by Rohit Parikh, and it studies the computational and logical analysis of social protocols and policies. It lies in the intersection of social choice theory and game theory, and is a subset of logic.

Such results have been presented in various talks including, *World Congress of Paraconsistency* in Kolkata and *Logic and the Foundations of Game and Decision Theory* in Bergen, and warmly received.

7. Partnerships and Cooperations

7.1. Regional Initiatives

7.1.1. SLAM: Schizophrenia and Language, Analysis and Modeling

Participants: Maxime Amblard [coordinator], Philippe de Groote, Sylvain Pogodalla, Karën Fort.

Schizophrenia is well-known among mental illnesses for the strength of the thought disorders it involves, and for their widespread and spectacular manifestations: from deviant social behavior to delusion, not to speak about affective and sensitive distortions. It aims at exploring a specific manifestation, namely disorders in conversational speech. This is an interdisciplinary research, both empirical and theoretical from several domains, namely psychology, philosophy, linguistic and computer science.

The SLAMproject starts for three years at the Maison des Sciences de l'Homme de Lorraine (MSH–Lorraine, USR 3261). While this year work was dedicated to the test protocol definition, the coming years will be devoted to building an open-access corpus of pathological uses of language.

The first transcriptions of pathological interviews are analyses. The management chain was implemented for disfluences and POS.

Other participants are: Denis Apotheloz (ATILF, Université de Lorraine),Valérie Aucouturier (Centre Léo Apostel, Université Libre de Bruxelles), Katarina Bartkova (ATILF, Université de Lorraine), Fethi Bretel (CHS Le Rouvray, Rouen), Michel Musiol (InterPSY, Université de Lorraine), Manuel Rebuschi (Archives Poincaré, Université de Lorraine).

The SLAM project was supported by the MSH–Lorraine, USR 3261, and won a PEPS project HuMaIn (mission pour l'interdisciplinarité du CNRS). The CNRS part of the budget allowed the organization of the second workshop which gather linguists, psychologists and computer scientists in december : http://discours.loria.fr

7.2. National Initiatives

7.2.1. ANR

7.2.1.1. Polymnie: Parsing and synthesis with abstract categorial grammars. From lexicon to discourse

Participants: Maxime Amblard, Philippe de Groote, Aleksandre Maskharashvili, Sylvain Pogodalla [coordinator], Sai Qian.

POLYMNIE[5] is a research project funded by the French national research agency (ANR). It relies on the grammatical framework of Abstract Categorial Grammars (ACG). A feature of this formalism is to provide the same mathematical perspective both on the surface forms and on the more abstract forms the latter correspond to. As a consequence:

- ACG allows for the encoding of a large variety of grammatical formalisms such as context-free grammars, Tree Adjoining grammars (TAG), etc.
- ACG defines two languages: an abstract language for the abstract forms, and an object language for the surface forms.

Importantly, the notions of object language and abstract language are relative to each other. If we can naturally see surface forms as strings for instance and abstract forms as the associated syntactic trees, we can also consider to associate this abstract form to a first order logical formula as surface (object) form. This property is central in our project as it offers a unified approach to text analysis and text generation, in particular considering the underlying algorithms and their complexity.

ACG definition uses type-theory and lambda-calculus. From this point of view, they smoothly integrate formal semantics models issuing from Montague's proposal. Theories that extend to the discourse level such as Discourse Representation Theory (DRT) and Dynamic Predicate Logic (DPL) were not initially formulated using lambda-calculus. But such formulations have been proposed. In particular, a formulation based on continuation semantics allows them to be expressed quite naturally in the ACG architecture. Dynamic effects of discourse, in particular those related to anaphora resultion or rhetorical relation inference, have then to be expressed by lexical semantics or computed from the syntactic rules as studied in the Inria Collaborative Research Project (ARC) CAuLD [6].

It has been shown that the discourse structure of texts plays a key role in their understanding. This is the case for both human readers and automatic processing systems. For instance, it can enhance text transformation systems such as the ones performing automatic summarization.

POLYMNIE focuses on studying and implementing the modelling of sentences and discourses in a compositional paradigm that takes into account their dynamics and their structures, both in parsing and in generation. To that end, we rely on the ACG framework. The kind of processing we are interested in relate to the automatic construction of summaries or to text simplification. This has to be considered in the limits of the modelling of the linguistic processes (as opposed to inferential processes for instance) these tasks involve.

The complexity of the phenomena, of their formal description, and of their interactions, require to set up a testing and development environment for linguistic modelling. It will consist in extending and stabilizing a software implementing the functionalities of the ACG framework. It will provide a tool for experimentation and validation of the approach.

Partners:

- Sémagramme people,
- Alpage (Paris 7 university & Inria Paris-Rocquencourt): Laurence Danlos (local coordinator), C. Braud, C. Roze, Éric Villemonte de la Clergerie,
- MELODI (IRIT, CNRS): Stergos Afantenos, Nicholas Asher (local coordinator), Juliette Conrath, Philippe Muller,
- Signes (LaBRI, CNRS): Jérôme Kirman, Richard Moot, Christian Retoré (local coordinator), Sylvain Salvati, Noémie-Fleur Sandillon-Rezer.

7.3. International Research Visitors

7.3.1. Visits of International Scientists

On the occasion of the workshop in honor of Hans KAMP we have invited in Nancy:

[5] http://semagramme.loria.fr/doku.php?id=projects:polymnie
[6] http://www.loria.fr/~pogodall/cauld/

- Nicholas ASHER, Université Paul Sabatier, France
- Paul DEKKER, Universiteit van Amsterdam, the Netherlands
- Bart GEURTS, Universiteit van Nijmegen, the Netherlands
- Irène HEIM, Massachusetts Institute of Technology, USA
- Klaus von HEUSINGER, Universität zu Köln, Germany
- Hans KAMP, Universität zu Stuttgart, Germany

7.3.2. Visits to International Teams

7.3.2.1. Sabbatical programme

Sylvain Pogodalla

Date: Aug 2014 - Jul 2015

Institution: Computational Linguistic Department of the University of Düsseldorf (Germany).

The objective of the research project deals with studying the syntax-semantics interface. It relies on two alternative approaches of this interface for mCSG: a unification based approach for Lexicalized Tree Adjoining Grammars (LTAG) [60], [61] as proposed in [57], [62], and a type-theoretic approach using Abstract Categorial Grammars (ACG) [80], [73], [74]. These two approaches provide the core mechanisms of structure mapping for the syntax-semantics interface. Because they both provide a perspective on the syntax-semantics interface for the same grammatical formalism, they offer an interesting meeting place for exchanges on the strength of each of the approaches. In the project, we focus on two of them: the role of lexical semantics and its interaction with the syntax-semantics design, and the integration of discourse related phenomena to the syntax-semantics interface. With that respect, the formal semantics expertise of the department in the modeling of tense and aspects plays is essential in enriching the approach.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific events organisation

Philippe de Groote, Maxime Amblard have organised a workshop on the occasion of the award of a Doctor Honoris Causa degree from the Université de Lorraine(hk-workshop.loria.fr).

Hans Kamp (Johan Anthony Willem Kamp) is a Dutch philosopher, linguist and logician. His work in formal semantics, most notably his Discourse Representation Theory, is widely used both in linguistics and in natural language processing. The workshop has focused on discourse semantics, especially on issues of its structure and dynamics, based on DRT.

It was held in the International Room of the MSH-Lorraine. The workshop was organised by Archives Poincaré, LORIA and ATILF, and it is supported by MSH-Lorraine, Université de Lorraine, CNRS and Inria.

Maxime Amblard organises the workshop (In)Cohérence du discours 2, with the SLAM project (Schizophrenia and Language: Analysis and Modelling).

The objective of the workshop was to discuss the latest advances in the modelling of discourses, in particular the kind held with pathological (e.g. schizophrenics). The adopted modelling paradigm is that of formal semantics, which falls within the scope of both linguistics and logic while also making ties to the philosophy of language.

8.1.1.1. Organizing committee membership

- Philippe de Groote and Maxime Amblard were members of the organisation committee of the workshop in honnor of Hans Kamp
- Maxime Amblard and Jiří Maršík were members of the organisation committee of the workshop (In)Coherence du discours 2

8.1.2. Member of the steering committee

- Philippe de Groote and Sylvain Pogodalla are members of the steering committe of the international conference series *Logical Aspects of Computational Linguistics (LACL)*.

8.1.2.1. Conference program committee membership

- Maxime Amblard:
  - member of the scientific committee of the conference *Traitement Automatique des Langues Naturelles (TALN)* 2014
  - member of the scientific committee of the workshop in honnor of Hans Kamp
  - member of the scientific committee of the workshop (In)Coherence du discours 2
- Philippe de Groote:
  - member of the program committee of the international conference *Formal Grammar (FG'14)*
  - member of the program committee of the international conference *Logical Aspects of Computational Linguistics (LACL'14)*
  - member of the program committee of the EACL 2014 workshop on *Type Theory and Natural Language Semantics (TTNLS)*
  - member of the scientific committee of the workshop in honnor of Hans Kamp
- Guy Perrier:
  - member of the program committee of the international conference PolTAL 2014, which was held in Warsaw, Poland.
  - member of the scientific committee of the conference *Traitement Automatique des Langues Naturelles (TALN) 2014*.
- Sylvain Pogodalla:
  - member of the scientific committee of the international conference *Traitement Automatique des Langues Naturelles (TALN) 2014*.
  - member of the scientific committee of international conference *Logical Aspects of Computational Linguistics (LACL) 2014*.

8.1.2.2. Reviewer

- Sylvain Pogodalla:
  - reviewer for the *International Conference on Computational Linguistics (Coling 2014)*.
  - reviewer for the international workshop on *Logic and Engineering of Natural Language Semantics 11 (LENLS11)*.

8.1.3. Journal

8.1.3.1. Editorial board membership

- Philippe de Groote:
  - area editor of the *FoLLI-LNCS series*.
  - associate editor of *Higher-Order and Symbolic Computation*.
  - member of the editorial board of *Cahiers du Centre de Logique*.

- Sylvain Pogodalla: member of the editorial board of the journal *Traitement Automatique des Langues*.

8.1.3.2. Reviewing activities

- Philippe de Groote: reviewer for the *Journal of Logic and Computation*, and for the conferences *LICS, LACL, and FG*.
- Sylvain Pogodalla: reviewer for the *Journal of Language Modelling*.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Master: Sylvain Pogodalla, *Computational Linguistic Formalisms: Semantics and Discourse*, 25h, M2, Université de Lorraine, France

Master: Jiří Maršík, *Fundamental Artificial Intelligence*, 10h, M2, Université de Lorraine, France

Master: Jiří Maršík, *Cognitif Aspect of Computational Linguistic*, 12h, M1, Université de Lorraine, France

Licence : Guy Perrier, *Initiation au TAL*, 20h, L2, Université de Lorraine, France

8.2.2. Supervision

PhD: Sai Qian, *Accessibility of Referents in Discourse Semantics*, Université de Lorraine, Nov. 7 2014, Philippe de Groote and Maxime Amblard

PhD in progress: Jiří Maršík, *Modeling Discourse in a Dynamics framework : formal integration and evaluation*, since september 2013, Philippe de Groote and Maxime Amblard

PhD in progress : Aleksandre Maskharashvili, *Generation and Discourse with Abstract Categorial Grammars*, since November 2012, Philippe de Groote and Sylvain Pogodalla.

8.2.3. Juries

- Philippe de Groote and Maxime Amblard were members of the PhD committee of Sai Qian, *Accessibility of Referents in Discourse Semantics*, Nov. 11 2014, Université de Lorraine.
- Philippe de Groote was president of the PhD committee of Hugo Férée, *Complexité d'ordre supérieur et analyse récursive*.
- Guy Perrier was president of the PhD committee of Shashi Narayan, *Generation and Simplification of Sentences*, Université de Lorraine, November 7 2014.
- Guy Perrier was member of the PhD committee of Arseniy Gorin, *Acoustic Model Structuring for Improving Automatic Speech Recognition Performance*, November 26 2014.
- Sylvain Pogodalla was member of the jury of the *Prix de thèse en Intelligence Artificielle* awarded by the Association Française pour l'Intelligence Artificielle (AFIA).

8.3. Popularization

- Maxime Amblard is member of the editorial board of interstice)i(, a french revue popularisation for computer sciences (http://interstices.info). He is the head of the rubric *informatique -ou presque- dans les films*.
- Maxime Amblard gave an interview for the *journal du CNRS* about Real Humans (a television drama series), may 2014.
- Maxime Amblard has participate to a television report about the SLAM project for *France 3 Lorraine*, may 2014.
- Maxime Amblard wrote an article about the SLAM project for the *Journal du CNRS*, january 2014.

- In November, Bruno Guillaume has demonstrated the prototype ZombiLingo and Jiří Maršík the ACGTK graphical output during the “Forum Sciences Cognitive” in Nancy.
- Jiří Maršík has participated to the regional final of the competition *Ma thèse en 180 secondes*.
- Jiří Maršík gave a talk to high school students at the Summer Camp of Mathematics and Physics (Letní matematicko-fyzikální soustředění) about treating words as fractions and algebraic expression where syntacticality is verified by cancelling out, Faculty of Mathematics and Physics of the Charles University (Matematicko-fyzikální fakulta Univerzity Karlovy).

8.4. Other Scientific Animation

- Maxime Amblard is an elected member of the “conseil d’UFR" of the UFR mathematics and computer sciences, until Dec. 2014.
- Maxime Amblard is an elected member of the “conseil de laboratoire" - Loria.
- Maxime Amblard is an elected member of the “conseil scientifique" of the Université de Lorraine (member of the board of the council since Nov. 2014).
- Maxime Amblard is an invited member of the “Pôle scientifique AM2I” of the Université de Lorraine.
- Maxime Amblard was member of the CA of the scientific association ATALA until june 2014.
- Maxime Amblard was member of the “comité de sélection" McF 4169 (CNU section 17 and 72) - Université Paris Descartes
- Bruno Guillaume is an elected member of the “Pôle scientifique AM2I” of the Université de Lorraine.
- Bruno Guillaume is a member of the Comipers (Inria commity for PhD and Post-doctoral selection).
- Bruno Guillaume was one of the two redactors of the CPER 2015-2020 project “Langues, Connaissances et Humanités Numériques” (Languages, Knowledge and Digital Humanities) in which ten laboratories of the Université de Lorrraine are implied.

9. Bibliography

Major publications by the team in recent years

[1] N. ASHER, S. POGODALLA. *A Montagovian Treatment of Modal Subordination*, in "20th Semantics and Linguistic Theory conference - SALT2010", Vancouver, Canada, N. LI, D. LUTZ (editors), 2011, http://hal.inria.fr/inria-00565616/en

[2] G. BONFANTE, B. GUILLAUME, M. MOREY, G. PERRIER. *Modular Graph Rewriting to Compute Semantics*, in "9th International Conference on Computational Semantics - IWCS 2011", Oxford, Royaume-Uni, J. BOS, S. PULMAN (editors), January 2011, pp. 65–74, http://hal.inria.fr/inria-00579244/en/

[3] G. BONFANTE, B. GUILLAUME, G. PERRIER. *Analyse syntaxique électrostatique*, in "Traitement Automatique des Langues (TAL)", 2003, vol. 44, n^{o} 3, pp. 93-120, Article dans revue scientifique avec comité de lecture nationale, http://hal.inria.fr/inria-00099582/en/

[4] P. DE GROOTE, M. KANAZAWA. *A Note on Intensionalization*, in "Journal of Logic, Language and Information", 2013, vol. 22, n^{o} 2, pp. 173-194 [*DOI :* 10.1007/s10849-013-9173-9], https://hal.inria.fr/hal-00909207

[5] B. GUILLAUME, G. PERRIER. *Interaction Grammars*, in "Research on Language & Computation", 2009, vol. 7, pp. 171–208

[6] P. DE GROOTE. *Towards abstract categorial grammars*, in "Association for Computational Linguistics, 39th Annual Meeting and 10th Conference of the European Chapter", Toulouse, France, none, July 2001, pp. 148-155, Colloque avec actes et comité de lecture. internationale, http://hal.inria.fr/inria-00100529/en

[7] P. DE GROOTE, M. KANAZAWA. *A Note on Intensionalization*, in "Journal of Logic, Language and Information", 2013, vol. 22, n° 2, pp. 173-194

[8] P. DE GROOTE, S. POGODALLA. *On the expressive power of abstract categorial grammars: Representing context-free formalisms*, in "Journal of Logic, Language and Information", 2004, vol. 13, n° 4, pp. 421–438

[9] P. DE GROOTE, S. POGODALLA, C. POLLARD. *About Parallel and Syntactocentric Formalisms: A Perspective from the Encoding of Convergent Grammar into Abstract Categorial Grammar*, in "Fundamenta Informaticae", 2011, vol. 106, n° 2-4, pp. 211-231 [*DOI :* 10.3233/FI-2011-384], http://hal.inria.fr/inria-00565598/en

[10] P. DE GROOTE. *Towards a Montagovian Account of Dynamics*, in "16th Semantics and Linguistic Theory conference - SALT2006", Tokyo, Japan, M. GIBSON, J. HOWELL (editors), 2006, http://elanguage.net/journals/index.php/salt/article/view/16.1

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] S. QIAN. *Accessibility of Referents in Discourse Semantics*, Université de Lorraine, November 2014, https://hal.inria.fr/tel-01104091

Articles in International Peer-Reviewed Journals

[12] M. AMBLARD, C. RETORÉ. *Partially Commutative Linear Logic and Lambek Caculus with Product: Natural Deduction, Normalisation, Subformula Property*, in "IFCoLog Journal of Logic and its Applications", June 2014, vol. 1, n° 1, pp. 53–94, 41 pages, https://hal.archives-ouvertes.fr/hal-01071642

International Conferences with Proceedings

[13] C. BASKENT. *Public Announcements, Topology and Paraconsistency*, in "Informal Proceedings of LOFT Conference 2014", Bregen, Norway, July 2014, https://hal.inria.fr/hal-01094782

[14] M. CANDITO, G. PERRIER, B. GUILLAUME, C. RIBEYRE, K. FORT, D. SEDDAH, É. VILLEMONTE DE LA CLERGERIE. *Deep Syntax Annotation of the Sequoia French Treebank*, in "International Conference on Language Resources and Evaluation (LREC)", Reykjavik, Iceland, May 2014, https://hal.inria.fr/hal-00969191

[15] A. COUILLAULT, K. FORT, G. ADDA, H. DE MAZANCOURT. *Evaluating Corpora Documentation with regards to the Ethics and Big Data Charter*, in "International Conference on Language Resources and Evaluation (LREC)", Reykjavik, Iceland, May 2014, https://hal.inria.fr/hal-00969180

[16] L. DANLOS, A. MASKHARASHVILI, S. POGODALLA. *An ACG Analysis of the G-TAG Generation Process*, in "INLG 2014 - 8th International Natural Language Generation Conference", Philadelphia, PA, United States, M. MITCHELL, K. MCCOY, D. MCDONALD, A. CAHILL (editors), Proceedings of the 8th International Natural Language Generation Conference (INLG), Association for Computational Linguistics, June 2014, pp. 35-44, https://hal.inria.fr/hal-00999595

[17] L. DANLOS, A. MASKHARASHVILI, S. POGODALLA. *An ACG View on G-TAG and Its g-Derivation*, in "LACL 2014 - Eight International Conference on Logical Aspects of Computational Linguistics", Toulouse, France, N. ASHER, S. SOLOVIEV (editors), Springer, June 2014, vol. 8535, pp. 70-82 [*DOI :* 10.1007/978-3-662-43742-1_6], https://hal.inria.fr/hal-00999633

[18] L. DANLOS, A. MASKHARASHVILI, S. POGODALLA. *Génération de textes : G-TAG revisité avec les Grammaires Catégorielles Abstraites*, in "TALN 2014 - 21ème conférence sur le Traitement Automatique des Langues Naturelles", Marseille, France, Actes de TALN 2014, Association pour le Traitement Automatique des Langues, July 2014, vol. 1, pp. 161-172, https://hal.inria.fr/hal-00999589

[19] P. DE GROOTE, Y. WINTER. *A type-logical account of quantification in event semantics*, in "Logic and Engineering of Natural Language Semantics 11", Tokyo, Japan, November 2014, https://hal.inria.fr/hal-01102261

[20] K. FORT, B. GUILLAUME, H. CHASTANT. *Creating Zombilingo, a Game With A Purpose for dependency syntax annotation*, in "Gamification for Information Retrieval (GamifIR'14) Workshop", Amsterdam, Netherlands, April 2014, https://hal.inria.fr/hal-00969157

[21] B. GUILLAUME, K. FORT, G. PERRIER, P. BEDARIDE. *Mapping the Lexique des Verbes du Français (Lexicon of French Verbs) to a NLP Lexicon using Examples*, in "International Conference on Language Resources and Evaluation (LREC)", Reykjavik, Iceland, May 2014, https://hal.inria.fr/hal-00969184

[22] M. LAFOURCADE, K. FORT. *Propa-L: a Semantic Filtering Service from a Lexical Network Created using Games With A Purpose*, in "International Conference on Language Resources and Evaluation (LREC)", Reykjavik, Iceland, May 2014, https://hal.inria.fr/hal-00969161

[23] J. MARŠÍK, M. AMBLARD. *Algebraic Effects and Handlers in Natural Language Interpretation*, in "Natural Language and Computer Science", Vienne, Austria, technical report containing the joint proceedings of the NLCS and NLSR workshops, July 2014, https://hal.archives-ouvertes.fr/hal-01079206

National Conferences with Proceedings

[24] G. ADDA, L. BESACIER, A. COUILLAULT, K. FORT, J. MARIANI, H. DE MAZANCOURT. *"Where the data are coming from?" Ethics, crowdsourcing and traceability for Big Data in Human Language Technology*, in "Crowdsourcing and human computation multidisciplinary workshop", Paris, France, CNRS, September 2014, https://hal.archives-ouvertes.fr/hal-01078045

[25] M. AMBLARD, K. FORT. *Étude quantitative des disfluences dans le discours de schizophrènes : automatiser pour limiter les biais*, in "TALN - Traitement Automatique des Langues Naturelles", Marseille, France, July 2014, pp. 292-303, https://hal.inria.fr/hal-01054391

[26] M. AMBLARD, K. FORT, M. MUSIOL, M. REBUSCHI. *L'impossibilité de l'anonymat dans le cadre de l'analyse du discours*, in "Journée ATALA éthique et TAL", Paris, France, November 2014, https://hal.archives-ouvertes.fr/hal-01079308

[27] K. FORT, B. GUILLAUME, V. STERN. *ZOMBILINGO : manger des têtes pour annoter en syntaxe de dépendances*, in "TALN - Traitement Automatique des Langues Naturelles", Marseille, France, July 2014, pp. 15-16, https://hal.inria.fr/hal-01054395

[28] G. PERRIER, M. CANDITO, B. GUILLAUME, C. RIBEYRE, K. FORT, D. SEDDAH. *Un schéma d'annotation en dépendances syntaxiques profondes pour le français*, in "TALN - Traitement Automatique des Langues Naturelles", Marseille, France, July 2014, pp. 574-579, https://hal.inria.fr/hal-01054407

Scientific Books (or Scientific Book chapters)

[29] M. AMBLARD, S. POGODALLA. *Modeling the Dynamic Effects of Discourse: Principles and Frameworks*, in "Dialogue, Rationality, and Formalism", M. REBUSCHI, M. BATT, G. HEINZMANN, F. LIHOREAU, M. MUSIOL, A. TROGNON (editors), Logic, Argumentation & Reasoning, Springer, 2014, vol. 3, pp. 247-282 [*DOI :* 10.1007/978-3-319-03044-9_12], https://hal.inria.fr/hal-00737765

[30] G. BONFANTE, B. GUILLAUME, M. MOREY, G. PERRIER. *Supertagging with Constraints*, in "Constraint and Language", P. BLACHE, H. CHRISTIANSEN, V. DAHL, D. DUCHIER, J. VILLADSEN (editors), Cambridge Scholar Publishing, 2014, https://hal.inria.fr/hal-01097999

[31] L. DANLOS, P. DE GROOTE, S. POGODALLA. *A Type-Theoretic Account of Neg-Raising Predicates in Tree Adjoining Grammars*, in "New Frontiers in Artificial Intelligence. JSAI-isAI 2013 Workshops, LENLS, JURISIN, MiMI, AAA, and DDS, Kanagawa, Japan, October 27-28, 2013, Revised Selected Papers", Y. NAKANO, K. SATOH, D. BEKKI (editors), Lecture Notes in Computer Science, Springer International Publishing, November 2014, vol. 8417, pp. 3-16 [*DOI :* 10.1007/978-3-319-10061-6_1], https://hal.inria.fr/hal-00868382

[32] V. DUFOUR-LUSSIER, B. GUILLAUME, G. PERRIER. *Parsing Coordination Extragrammatically*, in "Human Language Technology. Challenges for Computer Science and Linguistics. 5th Language and Technology Conference, LTC 2011, Poznan, Poland, November 25-27, 2011, Revised Selected Papers", Z. VETULANI, J. MARIANI (editors), Human Language Technology Challenges for Computer Science and Linguistics, Springer International Publishing, July 2014, 12 p. [*DOI :* 10.1007/978-3-319-08958-4_5], https://hal.inria.fr/hal-00921033

[33] K. FORT, G. ADDA, B. SAGOT, J. MARIANI, A. COUILLAULT. *Crowdsourcing for Language Resource Development: Criticisms About Amazon Mechanical Turk Overpowering Use*, in "Human Language Technology Challenges for Computer Science and Linguistics", Z. VETULANI, J. MARIANI (editors), Springer International Publishing, July 2014, pp. 303-314 [*DOI :* 10.1007/978-3-319-08958-4_25], https://hal.inria.fr/hal-01053047

[34] C. GARDENT, Y. PARMENTIER, G. PERRIER, S. SCHMITZ. *Lexical Disambiguation in LTAG using Left Context*, in "Human Language Technology. Challenges for Computer Science and Linguistics. 5th Language and Technology Conference, LTC 2011, Poznan, Poland, November 25-27, 2011, Revised Selected Papers", Z. VETULANI, J. MARIANI (editors), Lecture Notes in Computer Science (LNCS) series / Lecture Notes in Artificial Intelligence (LNAI) subseries, Springer, July 2014, vol. 8387, pp. 67-79 [*DOI :* 10.1007/978-3-319-08958-4_6], https://hal.inria.fr/hal-00921246

[35] M. REBUSCHI, M. AMBLARD, M. MUSIOL. *Using SDRT to analyze pathological conversations. Logicality, rationality and pragmatic deviances*, in "Interdisciplinary Works in Logic, Epistemology, Psychology and Linguistics: Dialogue, Rationality, and Formalism", M. REBUSCHI, M. BATT, G. HEINZMANN, F. LIHOREAU, M. MUSIOL, A. TROGNON (editors), Logic, Argumentation & Reasoning, Springer, 2014, pp. 1–24, https://hal.archives-ouvertes.fr/hal-00910725

Research Reports

[36] G. PERRIER. *FRIGRAM: a French Interaction Grammar*, Inria Nancy, December 2014, n° RR-8323, 257 p. , https://hal.inria.fr/hal-00840254

Other Publications

[37] C. BASKENT. *Some Non-Classical Approaches to the Branderburger-Keisler Paradox*, December 2014, https://hal.inria.fr/hal-01094784

[38] C. BASKENT. *Topological Semantics for da Costa Paraconsistent Logics Cω and C*ω*, 2014, https://hal.inria.fr/hal-01094786

[39] C. BASKENT. *Which Society, Which Software?*, 2014, https://hal.inria.fr/hal-01094785

References in notes

[40] Y. BAR-HILLEL. *Out of the pragmatic wastebasket*, in "Linguistic Inquiry", 1971, vol. 2, n° 3, pp. 401–407

[41] C. BARKER. *Continuations and the Nature of Quantification*, in "Natural Language Semantics", 2002, vol. 10, n° 3, pp. 211–242

[42] C. BARKER. *Continuations in Natural Language*, in "Proceedings of the Fourth ACM SIGPLAN Continuations Workshop (CW'04)", 2004

[43] C. BASKENT. *Game Semantics: A Paraconsistent Approach*, under review

[44] C. BASKENT. *Inquiry, Refutations and the Inconsistent*, 2015, https://hal.inria.fr/hal-01094783

[45] G. BONFANTE, B. GUILLAUME, M. MOREY. *Dependency Constraints for Lexical Disambiguation*, in "11th International Conference on Parsing Technologies - IWPT'09", Paris, France, Association for Computational Linguistics, October 2009, pp. 242-253, https://hal.inria.fr/inria-00440795

[46] G. BONFANTE, B. GUILLAUME, G. PERRIER. *Polarization and abstraction of grammatical formalisms as methods for lexical disambiguation*, in "20th Conference on Computational Linguistics - CoLing'2004", Genève, Suisse, 2004, pp. 303-309, Colloque avec actes et comité de lecture internationale, https://hal.inria.fr/inria-00099886

[47] M. CANDITO, D. SEDDAH. *Le corpus Sequoia : annotation syntaxique et exploitation pour l'adaptation d'analyseur par pont lexical*, in "Proc. of TALN", Grenoble, France, June 2012

[48] N. CHOMSKY. *Three models for the description of language*, in "Information Theory, IRE Transactions on", 1956, vol. 2, n° 3, pp. 113–124

[49] A. CHURCH. *A formulation of the simple theory of types*, in "The journal of symbolic logic", 1940, vol. 5, n° 2, pp. 56–68

[50] L. DANLOS. *D-STAG : un formalisme d'analyse automatique de discours basé sur les TAG synchrones*, in "T.A.L.", 2009, vol. 50, n° 1, pp. 111–143, http://hal.inria.fr/inria-00524743/en/

[51] L. DANLOS. *D-STAG: a Formalism for Discourse Analysis based on SDRT and using Synchronous TAG*, in "14th conference on Formal Grammar - FG 2009", P. DE GROOTE, M. EGG, L. KALLMEYER (editors), LNCS/LNAI, Springer, 2011, vol. 5591, pp. 64–84, http://dx.doi.org/10.1007/978-3-642-20169-1_5

[52] L. DANLOS. *G-TAG : Un formalisme lexicalisé pour la génération de textes inspiré de TAG*, in "Traitement Automatique des Langues", 1998, vol. 39, n° 2

[53] J. DUBOIS, F. DUBOIS-CHARLIER. *Les verbes français*, Larousse-BordasParis, France, 1997

[54] M. FELLEISEN, D. P. FRIEDMAN, E. KOHLBECKER, B. DUBA. *A syntactic theory of sequential control*, in "Theor. Comput. Sci.", June 1987, vol. 52, pp. 205–237, http://dx.doi.org/10.1016/0304-3975(87)90109-5

[55] M. FELLEISEN, R. HIEB. *The revised report on the syntactic theories of sequential control and state*, in "Theoretical Computer Science", 1992, vol. 103, n° 2, pp. 235 - 271 [*DOI :* 10.1016/0304-3975(92)90014-7], http://www.sciencedirect.com/science/article/pii/0304397592900147

[56] D. GALLIN. *Intensional and higher-order modal logic: with applications to Montague semantics*, North-Holland mathematics studies, North-Holland, 1975, http://books.google.fr/books?id=jkiNo5Rok7cC

[57] C. GARDENT, L. KALLMEYER. *Semantic construction in Feature-Based TAG*, in "Proceedings of the 10th Meeting of the European Chapter of the Association for Computational Linguistics (EACL)", 2003, pp. 123–130, http://www.aclweb.org/anthology/E03-1030

[58] J. GROENENDIJK, M. STOKHOF. *Dynamic predicate logic*, in "Linguistics and philosophy", 1991, vol. 14, n° 1, pp. 39–100

[59] B. GUILLAUME, G. PERRIER. *Interaction Grammars*, in "Research on Language & Computation", 2009, vol. 7, pp. 171–208

[60] A. K. JOSHI, L. S. LEVY, M. TAKAHASHI. *Tree Adjunct Grammars*, in "Journal of Computer and System Sciences", 1975, vol. 10, n° 1, pp. 136–163, http://dx.doi.org/10.1016/S0022-0000(75)80019-5

[61] A. K. JOSHI, Y. SCHABES. *Tree-adjoining grammars*, in "Handbook of formal languages", G. ROZENBERG, A. K. SALOMAA (editors), Springer, 1997, vol. 3, chap. 2

[62] L. KALLMEYER, M. ROMERO. *Scope and Situation Binding for LTAG*, in "Research on Language and Computation", 2008, vol. 6, n° 1, pp. 3–52, http://dx.doi.org/10.1007/s11168-008-9046-6

[63] H. KAMP, U. REYLE. *From discourse to logic: Introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory*, Kluwer Academic Dordrecht, The Netherlands, 1993, vol. 42

[64] M. KANAZAWA, S. SALVATI. *Generating Control Languages with Abstract Categorial Grammars*, in "Proceedings of the 12th conference on Formal Grammar", 2007

[65] J. LAMBEK. *The mathematics of sentence structure*, in "The American Mathematical Monthly", 1958, vol. 65, n^{o} 3, pp. 154–170

[66] F. MEUNIER. *Implantation du formalisme de génération G-TAG*, Université Paris 7 — Denis Diderot, 1997

[67] R. MONTAGUE. *English as a formal language*, in "Linguaggi nella societae nella tecnica", 1970, pp. 189–224

[68] R. MONTAGUE. *Universal grammar*, in "Theoria", 1970, vol. 36, n^{o} 3, pp. 373–398

[69] R. MONTAGUE. *The Proper Treatment of Quantification in Ordinary English*, in "Formal Semantics", 1973, pp. 221–242

[70] M. MOORTGAT. *Categorial type logics*, in "Handbook of logic and language", 1997, pp. 93–177

[71] C. MORRIS. *Foundations of the Theory of Signs*, University of Chicago Press, 1938, vol. 1

[72] G. PERRIER. *A French Interaction Grammar*, in "International Conference on Recent Advances in Natural Language Processing - RANLP 2007", Borovets, Bulgarie, G. ANGELOVA, K. BONTCHEVA, R. MITKOV, N. NICOLOV, K. SIMOV (editors), INCOMA Ltd, Shoumen, Bulgaria, 2007, pp. 463-467, http://hal.inria.fr/inria-00184108/en/

[73] S. POGODALLA. *Computing Semantic Representation: Towards ACG Abstract Terms as Derivation Trees*, in "Seventh International Workshop on Tree Adjoining Grammar and Related Formalisms - TAG+7", Vancouver, BC, Canada, 2004, pp. 64-71, Colloque avec actes et comité de lecture. Internationale, https://hal.inria.fr/inria-00107768

[74] S. POGODALLA. *Advances in Abstract Categorial Grammars: Language Theory and Linguistic Modeling. ESSLLI 2009 Lecture Notes, Part II*, 2009, https://hal.inria.fr/hal-00749297

[75] S. QIAN. *Accessibility of Referents in Discourse Semantics*, Université de Lorraine, November 2014, https://hal.inria.fr/tel-01104091

[76] C. RETORÉ, S. SALVATI. *A Faithful Representation of Non-Associative Lambek Grammars in Abstract Categorial Grammars*, in "Journal of Logic Language and Information", April 2010, vol. 19, n^{o} 2, pp. 185–200, http://hal.inria.fr/inria-00409557/en/

[77] J. ROGERS. *A model-theoretic framework for theories of syntax*, in "Proceedings of the 34th annual meeting on Association for Computational Linguistics", Association for Computational Linguistics, 1996, pp. 10–16

[78] C.-C. SHAN. *Delimited continuations in natural language: quantification and polarity sensitivity*, in "CoRR", 2004, vol. cs.CL/0404006

[79] C. STRACHEY, C. P. WADSWORTH. *Continuations: A Mathematical Semantics for Handling Full Jumps*, in "Higher-Order and Symbolic Computation", 1974, vol. 13, pp. 135–152

[80] P. DE GROOTE. *Towards abstract categorial grammars*, in "Association for Computational Linguistics, 39th Annual Meeting and 10th Conference of the European Chapter", Toulouse, France, July 2001, pp. 148-155, Colloque avec actes et comité de lecture. internationale, https://hal.inria.fr/inria-00100529

[81] P. DE GROOTE. *Towards abstract categorial grammars*, in "Proceedings of the 39th Annual Meeting on Association for Computational Linguistics", Association for Computational Linguistics, 2001, pp. 252–259

[82] P. DE GROOTE. *Tree-Adjoining Grammars as Abstract Categorial Grammars*, in "TAG+6, Proceedings of the sixth International Workshop on Tree Adjoining Grammars and Related Frameworks", Università di Venezia, 2002, pp. 145–150

[83] P. DE GROOTE. *Towards a montagovian account of dynamics*, in "Semantics and Linguistic Theory (SALT) 16", M. GIBSON, J. HOWELL (editors), 2006, pp. 1-16

[84] P. DE GROOTE, S. POGODALLA. *m-linear context-free rewriting systems as abstract categorial grammars*, in "oceedings of Mathematics of Language - MOL-8", 2003

[85] P. DE GROOTE, S. POGODALLA. *On the expressive power of abstract categorial grammars: Representing context-free formalisms*, in "Journal of Logic, Language and Information", 2004, vol. 13, n^{o} 4, pp. 421–438

[86] J. VAN BENTHEM. *Essays in logical semantics*, Reidel, 1986, n^{o} 29

[87] K. VAN DEN EYNDE, C. BLANCHE-BENVENISTE. *Syntaxe et mécanismes : présentation de l'approche pronominale*, in "Cahiers de Lexicologie", 1978, vol. 32, pp. 3–37