Discursive analysis of itineraries in a historical regional corpus of travels

syntax, semantics & pragmatics in a unified type theoretical framework

WITH A CLOSE-UP ON SYSTEM F

Richard Moot[Bx], Laurent Prévot[Aix], Christian Retoré[Bx]

Bx: LaBRI, INRIA, Université de Bordeaux

Aix: LPL Université de Provence

ITIPY meeting, Toulouse, 2011 Nov. 7

Providing a type theoretical frame work for a treatment from syntax to discourse (CG syntax, compositional DRT semantics, meaning transfers, discourse relations)

1. corpus and objective,
   the virtual traveller problem
2. categorial parser with DRS outputs
3. extending the type system for lexical pragmatics
4. a lexicon and an analysis
   involving a virtual traveller

No interpretation of the semantic representation.
Faithful modelling of the linguistic analyses of motion verbs done by others.

Part I

Data, question and outline

1. A case study, a field for semantic experiments

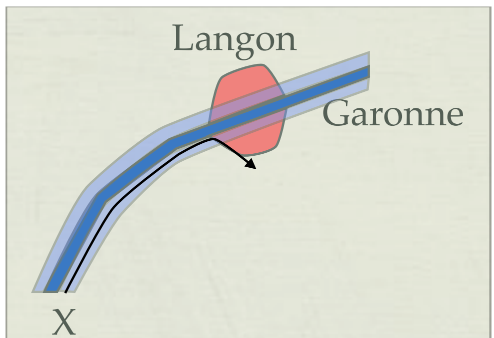
Corpus: French, XVII-XX century (mainly XIX), travel stories through the Pyrenees (576.334 words).

Goal: given some text, can we reconstruct the itinerary followed by the traveller?

More concretely: we want to transform the text into some semantic representation which allows us to infer the itinerary.

2. Example

(1) Jusqu'à Langon, nous avons longé la Garonne, traversant un véritable jardin rempli de vignes et d'arbres fruitiers.
Until Langon, we went along the Garonne, making our way through what seemed like a garden filled with vines and fruit trees.



3. From parsing to semantics within type theory

- Multimodal categorial grammar
  - syntactic lexicon acquired from corpora
    Categorial grammar
  - semantic/lexicon hand written (smaller prototype)
    lambda DRT
- lexicon designed for semantics and pragmatics
  meaning transfer functions, second order lambdas
- ? finding out narration, elaboration, background relations
  SDRT
- ?? itinerary reconstruction
  spatio-temporal reasoning using our geographic data base, tense/aspect, etc.

To facilitate this process, we advocate the use of variable types (Girard's system F).

4. A particular phenomenon: “fictive motion”

For though it be lawful to say, for example, in common speech, the way goeth, or leadeth hither, or thither; the proverb says this or that (whereas ways cannot go, nor proverbs speak); yet in reckoning, and seeking of truth, such speeches are not to be admitted.
Hobbes, Leviathan (1651), Chapter V, Of Reason and Science.

Many of the authors of our corpus don’t follow this advice. In the cognitive literature (Talmy), the following examples are called “fictive movement”.

(2) The path descended abruptly.

(3) The road runs along the coast for two hours.

(4) The fence zigzags from the plateau to the valley.

(5) The highway crawls through the city.

This does not mean that someone actually follows the path.
How do we model this?

5. Examples from our corpus

(6) Nous coupons ici un sentier qui vient du port de Barroude (...)
Here, we cross a path which comes from the pass of Barroude

(7) Plus loin, de nobles hêtres montent sur le versant (...)
Further away, noble beeches climb the slope

(8) (...) cette route qui monte sans cesse pendant deux lieues
this road which climbs incessantly for two miles

(9) Le chemin pavé de calcaire et de pierres luisantes (...) serpente à travers fourrés de buis et de noisetiers
The road paved with limestone and shining stones winds across buxus and hazels shrubbery

(10) Puis, cinq minutes nous conduisent à un petit pont (...) qui nous porte sur la rive droite.
Afterwards, five minutes take us to a small bridge ... which carries us to the right bank.

Part II

Categorial syntax and semantics

6. Grail categorial parser and French grammar

- Grail is a general-purpose parser for (multimodal) categorial grammars.
- A wide-coverage French grammar has been semi-automatically extracted from the French Treebank.
- On the basis of the 382.145 words and 12.822 sentence of the treebank, the extraction algorithm extracts 883 different formulas, of which 664 occur more than once.
- Many frequent words are assigned *many* different formulas.
- Standard statistical methods (*supertagging*) help with lexical disambiguation.

7. Number of entries for common words and POS tags

et	CONJ	71
,	PONCT	62
à	PRP	55
plus	ADV	44
ou	CONJ	42
est	VERB	39
être	INF	36
en	PRP	34
a	VERB	31

ADV	206
VERB	175
PRP	149
CONJ	92
PONCT	89

8. Syntactic categories

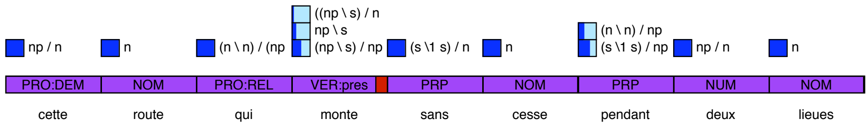
Atomic categories: *np* (noun phrase), *n* (common noun), *s* (sentence) and *pp* (prepositional phrase) (the actual set of categories is slightly more detailed).
Categories:

- atomic categories
- A/B whenever A and B are categories
- $B\backslash A$ whenever A and B are categories

Rules:

$$\frac{A/B \quad B}{A}\ /E \qquad\qquad \frac{\begin{array}{c}\ldots[B]\\ \vdots\\ A\end{array}}{A/B}\ /I$$

9. Example



- The supertagger assigns supertag with a confidence level (indicated by the darker blue part of the square before the formula)
- when it is less sure it proposes more alternatives
- Most words only have a single formula is proposed by the supertagger, the verb “monte” (which does not occur at all in the training corpus) has three possible formulas and the preposition “pendant” has two.

10. Semantics

- Categorial derivations (being a proper subset of derivations in multiplicative intuitionistic linear logic) correspond to (simply typed) linear lambda terms.
- This makes the connection to Montague grammar particularly transparent.

$$\frac{A/B : f^{U \rightarrow T} \quad B : x^{U}}{A : (f\ x)^{T}} \ /E$$

$$\begin{array}{c} \ldots [B : x^{U}] \\ \vdots \\ \dfrac{A : t^{T}}{A/B : \lambda x^{U} t} \ /I \end{array}$$

11. Language, Metaphysics and Ontology

Metaphysics

- What is there?
- What kind of things are there and how are they related?

Natural Language Metaphysics Bach (1986)

- What do people talk as if there is?
- What kinds of things and relations along them does one need in order to exhibit the structural of meanings that natural languages seem to have?

But...
"We have learned to be weary, however, of what the surface of language suggests, especially when it comes to ontology" Davidson (1970)

12. Different sorts

Montague: one type of entity e.

"although there may be real metaphysical doubts whether the Great Pyramid and the Battle of Waterloo are fundamentally different in kind, natural language seems to advise us that they are different" Bach (1986)

Many syntactically well-formed sentences are semantically ill-formed in a way which is more easily and better modeled by using different sorts (Ty_n, Gallin Muskens).

(11) The Battle of Waterloo took place on Sunday 18 June 1815.

(12) # The Great Pyramid took place on Sunday 18 June 1815.

If we distinguish between events and physical objects we can say that sentence 12 is nonsense (instead of simply false).

12.1. Events

Davidson (1967) looks at sentences such as the following

(13) Brutus stabbed Caesar with a knife on the Ides of March.

(14) The stabbing of Ceasar (with a knife) by Brutus took place on the Ides of March.

and proposes to analyse them as follows (where e is a variable denoting an event)

$$\exists e.\mathit{stab}(e, \mathit{Brutus}, \mathit{Ceasar}) \wedge \mathit{with}(e, \mathit{knife}) \wedge \mathit{date}(e, \mathit{IdesofMarch})$$

This, as he notes, has the pleasant consequence that the following formulas are all *derivable* from this semantic representation.

$$\vdash \exists e.\mathit{stab}(e, \mathit{Brutus}, \mathit{Ceasar}) \wedge \mathit{with}(e, \mathit{knife})$$
$$\vdash \exists e.\mathit{stab}(e, \mathit{Brutus}, \mathit{Ceasar}) \wedge \mathit{date}(e, \mathit{IdesofMarch})$$
$$\vdash \exists e.\mathit{stab}(e, \mathit{Brutus}, \mathit{Ceasar})$$

12.2. Other sorts

Besides events/eventualities, we use several other types of entities including:

- locations
- paths
- physical objects
- information

It’s a flat ontology (but we’ll come back to that)

12.3. No Entity Without Identity...

Quine (1969) famously said “no entity without identity”.

How do we decide whether or not two events are identical?

On the other hand, the ship of Theseus (as reported by Plutarch) and the river of Heraclitus show that identity for physical objects is not so easy either.

"We must not try to resolve the metaphysical questions first, and then construct a meaning-theory in the light of the answers. We should investigate how our language actually functions, and how we can construct a workable systematic description of how it functions; the answers to those questions will then determine the answers to the metaphysical ones.”
Dummet (1991, p. 338)

13. Semantics with lambda-DRT

We use a λ-DRT entities/individuals divided into several sorts Ty_n.
Extension to TY_n without difficulty nor suprise: e can be divided in several kind of entities.
Let us show an example of a lexicon, and the treatment of an example.
$\oplus$ is the DRS merge operation.
All movement verbs are analysed with a three argument predicate *travel*(e, x, p), where e is a movement event, x is the moving physical entity (typically a person) and p is a path.

word phrase syntactic type	lambda-DRS
descend $np\backslash s$	$\lambda x^{person}\lambda e^{event}\begin{array}{\|l\|}\hline p^{path} \\ \hline \mathit{travel}(e,x,p) \\ \mathit{height}(\mathit{source}(p)) \\ \quad > \mathit{height}(\mathit{destination}(p)) \\ \hline \end{array}$
Jean $s/(np\backslash s)$	$\lambda P^{person\rightarrow event\rightarrow t}\lambda e^{event}\begin{array}{\|l\|}\hline y^{person} \\ \hline \mathit{Jean}(y) \\ \hline \end{array} \oplus ((P\ y)\ e)$

14. Syntactic / semantic analysis

Syntactic analysis is proving
Jean descend $\vdash s$ ....
and it tells us that the semantic is the application of the semantic λ-term of *Jean* to the one of *descend*

$$\left((\lambda P^{person \to event \to t} \lambda e^{event} \begin{array}{|c|}\hline y^{person} \\ \hline Jean(y) \\ \hline \end{array} \oplus ((P\ y)\ e)) \right.$$

$$\left. \lambda x^{person} \lambda e^{event} \begin{array}{|l|}\hline p^{path} \\ \hline travel(e, x, p) \\ height(source(p)) \\ \quad > height(destination(p)) \\ \hline \end{array} \right)$$

$$\to_\beta$$

$$\lambda e^{event} \begin{array}{|l|}\hline y^{person} \\ \hline \mathit{Jean}(y) \\ \hline \end{array} \oplus \left((\lambda x^{person} \lambda e^{event} \begin{array}{|l|}\hline p^{path} \\ \hline \mathit{travel}(e, x, p) \\ \mathit{height}(\mathit{source}(p)) \\ \qquad > \mathit{height}(\mathit{destination}(p)) \\ \hline \end{array} y \right)$$

$$\rightarrow_\beta$$

$$\lambda e^{event} \begin{array}{|l|}\hline y^{person} \\ \hline \mathit{Jean}(y) \\ \hline \end{array} \oplus ((\lambda e^{event} \begin{array}{|l|}\hline p^{path} \\ \hline \mathit{travel}(e, y, p) \\ \mathit{height}(\mathit{source}(p)) \\ \qquad > \mathit{height}(\mathit{destination}(p)) \\ \hline \end{array}) \, e)$$

$$\rightarrow_\beta$$

$$\lambda e^{event} \begin{array}{|l|}\hline y^{person} \\ \hline \textit{Jean}(y) \\ \hline \end{array} \oplus \begin{array}{|l|}\hline p^{path} \\ \hline \textit{travel}(e, y, p) \\ \textit{height}(\textit{source}(p)) \\ \quad > \textit{height}(\textit{destination}(p)) \\ \hline \end{array}$$

$\longrightarrow$ DRS merge

$$\lambda e^{event} \begin{array}{|l|}\hline y^{person} \; p^{path} \\ \hline \textit{Jean}(y) \\ \textit{travel}(e, y, p) \\ \textit{height}(\textit{source}(p)) > \textit{height}(\textit{destination}(p)) \\ \hline \end{array}$$

$\longrightarrow$ Existential closure

$e^{event}\ y^{person}\ p^{path}$
$\mathit{Jean}(y)$
$\mathit{travel}(e, y, p)$
$\mathit{height}(\mathit{source}(p))$
$\quad > \mathit{height}(\mathit{destination}(p))$

Part III
Extending the type system: ΛTy_n, $F-DRT$

15. Second order types (Girard's F).

Ty_n (several base types) filters the sort of the argument according to lexical constraints, but....

varying types, flexible types are useful

and uniform operations on types as well.

Such features exists in Girard system F and can be used for

- co-predication,
- (generalized) quantification,
- plurals

and all operation that act uniformly upon all types, one can also add type variables and quantification over types.

16. More general types and terms. Second order types (Girard's F).

- Constants e (or $e_1, \ldots, e_n$ in a multisorted system) and t, as well as any type variable α in P, are types.
- Whenever T is a type and α a type variable which may but need not occur in T, $\Pi.\alpha.\ T$ is a type.
- Whenever T_1 and T_2 are types, $T_1 \to T_2$ is also a type.

17. More general types and terms. Second order terms (Girard's F).

- A variable of type T i.e. $x : T$ or x^T is a *term*. Countably many variables of each type.
- $(f\ \tau)$ is a term of type U whenever $\tau : T$ and $f : T \to U$.
- $\lambda x^T.\ \tau$ is a term of type $T \to U$ whenever $x : T$, and $\tau : U$.
- $\tau\{U\}$ is a term of type $T[U/\alpha]$ whenever $\tau : \Lambda\alpha.\ T$, and U is a type.
- $\Lambda\alpha.\tau$ is a term of type $\Pi\alpha.T$ whenever α is a type variable, and $\tau : T$ without any free occurrence of the type variable α.

18. More general types and terms. Second order reduction.

The reduction is defined as follows:

- $(\Lambda\alpha.\tau)\{U\}$ reduces to $\tau[U/\alpha]$ (remember that α and U are types).
- $(\lambda x.\tau)u$ reduces to $\tau[u/x]$ (usual reduction).

Reduction is strongly normalising and confluent (Girard, 1971): every term of every type admits a unique normal form which is reached no matter how one proceeds.

F can compute all recursive functions whose totality can be proved in second order Peano arithmetic (classical $\sim$ intutionistic for such statements).

Example of a function that cannot be computed: normalisation of the terms of system F .

19. More general types and terms. A second order example.

Given two predicates $P^{\alpha\to t}$ and $Q^{\beta\to t}$
over entities of respective kinds α and β
when we have two morphisms from ξ to α and to β
we can coordinate entities of type ξ:
$\Lambda\xi\lambda x^{\xi}\lambda f^{\xi\to a}\lambda g^{\xi\to b}.(and\ (P\ (f\ x))(Q\ (g\ x)))$

One can even quantify over the predicates P, Q and the types α, β to which they apply:
$\Lambda\alpha\Lambda\beta\lambda P^{\alpha\to t}\lambda Q^{\beta\to t}\Lambda\xi\lambda x^{\xi}\lambda f^{\xi\to\alpha}\lambda g^{\xi\to\beta}.(\mathrm{and}\ (P\ (f\ x))(Q\ (g\ x)))$

20. Apology for system F : general remarks

- used for the syntax of semantics (a.k.a. metalogic, glue logic)
- the formulae of semantics are the usual ones
- a single constant, e.g. for the quantifier $\forall$ or the choice function ι which is specialized for each type
- less types (constrained) than formulae with a free variable (e.g. types $\sim$ comparison classes).

It is also the type system of the polymorphic functional programming languages ML, CaML,...

21. Apology for system F : expressive power and simplicity

Universal quantification provides (internally) inductive data types (as in ML or CaML):

- cartesian product $A \times B \equiv \Pi X.(A \rightarrow (B \rightarrow X) \rightarrow X$
- existential quantification over types $\exists p.\ A[p] \equiv \Pi q.\ (\Pi p.\ (A[p] \rightarrow q)) \rightarrow q$
- booleans $\Pi X.X \rightarrow X \rightarrow X$
- integers $\Pi X.(X \rightarrow X) \rightarrow (X \rightarrow X)$
- lists of objects of type α: $\Pi X.X \rightarrow (\alpha \rightarrow X \rightarrow X) \rightarrow X$

A way to characterise is expressive power is to say that in this system one can define exactly all the total recursive functions whose totality can be proved in second order Peano arithmetic (classical or intuitionistic it makes no difference).
Only 4 rules (introduction and elimination of λ and Λ) and 2 similar β-reduction patterns.

22. Apology for system F : a glue logic for your favourite logic

The glue logic $\neq$ the logic used for formulating semantics

The system F (intuitionistic *second* order propositional logic) can glue *n*-order logic formulae $n \in \mathbb{N} \cup \{\omega\}$ (remember: simply typed λ-calculus with $\mathbf{e}$ and $\mathbf{t}$ glues ω order logic)

When λ-calculus constants are the ones of n order logic, the normal λ-terms of F of type t are formulae of n order logic.

System F is widely relevant for the logical syntax of semantics:

- lexical pragmatics (meaning transfers, dot objects)
- (generalised) quantification
- plurals

23. Apology for system F: subtypes?

Subtyping is not compatible with system F despite some attempts by Cardelli et al.

But subtyping is not the IS A relation that we are looking for.

Subtyping: inclusions between complex types like $a \rightarrow b$ are all the ones derived from inclusions on a and inclusions on b.

Does subtyping on verb types derives from subtyping of its arguments, subject, object, etc. ?

Does classifications of "food"and "eaters" provide a classification of "eating" verbs (*swallow, taste, appreciate*)?

Worse:
- language does not allow all the ontological inclusions?
- does idiosyncratic linguistic inclusions define an order?

24. Apology for system F : is F as safe as TT alternatives?

Formal complexity

- Algorithmic complexity is not an issue (syntax performs parsing, semantics β-reduces simple terms)
- Martin-Löf TT (used by Z. Luo) many rules, many variants
- F defined by 4 rules and 2 reduction patterns

Coherence proofs use the axiom of comprehension CA (which defines a set $\{X|P(X)\}$ from a formula P)

- TT is simpler comprehension for $P \in \Pi^1_1$
- F more complicated: any P

Records are in both system (products with named projections)
Dependent types: most TT have them, and not F but they can be added to F if one wishes to (constructions = F $_\omega$+dependent types) A question: why do we need dependent types?

Part IV

The theoretical lexicon at work: the virtual traveller

25. Dot objects

Several words can have multiple incompatible sorts. The most famous example of Pustejovsky (1995) is the book, which is both a physical object (and therefore can be picked up and be heavy) and information (and therefore something which can be interesting or easy to read).
Examples like the following

(15) This book is heavy but interesting.

argue against a naive “lexical ambiguity” approach: which would incorrectly analyse sentence 15 as incoherent.

$$\lambda x^{p} (\mathit{book}\ x)$$
$$\lambda x^{i} (\mathit{book}\ x)$$
$$\lambda P^{p \to t} \lambda x^{p} (\mathit{heavy}\ x) \wedge (P\ x)$$
$$\lambda P^{i \to t} \lambda x^{i} (\mathit{interesting}\ x) \wedge (P\ x)$$

26. A Type-Theoretic Solution

A more appropriate solution is to say that there is a functional correspondance between the book as a physical object and its information content — this seems quite reasonable: children learn this function at school, and OCR software goes a long way towards implementing this function.

$$\lambda x^{p} (\mathit{book}\ x)$$

$$\lambda P^{p \rightarrow t} \lambda x^{i} ((f^{(p \rightarrow t) \rightarrow (i \rightarrow t)} P)\ x)$$

This is essentially the solution of Bassac, Mery and Retoré (2010) and Mery (2011): a lexical entry can specify some optional transformations which change the sortal information of the lambda term assigned to a word.

27. More dot-objects and copredications

See Asher (2011) for discussion and many more examples.

(16) Lunch was delicious but took forever. (food + event)

(17) I saw the Collosseum in my tourist guide and wanted to go there. (anaphor: artifact + place)

(18) The road was built in 1825 and runs from the Regent's residence at Carlton house to All Souls Church. (artifact + path)

Other examples are less good.

(19) ? Barcelona organized the 1992 olympics and won the 2010-2011 Champions League (local government + football club)

(20) # Barcelona is the capital of Catalunya and won the 2010-2011 Champions League.

(21) # Washington borders the Potomac and attacked Iraq.

28. Principles of our lexicon

- Remains within realm of Montagovian compositional semantics (though possibly without possible worlds semantics).
- Allows both predicate and argument to contribute lexical information to the compound.
- Works with λ-DRT.

We advocate a system based on *optional modifiers* which can account for lexical idiosyncrasies.
Second-order typing, like Girard's F system is needed for arbitrary modifiers:

$$\Lambda\alpha\lambda x^{A} y^{\alpha} f^{\alpha\to R}.((\mathsf{read}^{A\to R\to t}\ x)\ (f\ y))$$

29. A lexical entry

- syntactic category x
- A standard Λ-term of type x^* attached to the main sense:
  - Used for compositional purposes
  - Comprising detailed typing information
  - Including slots for optional modifiers
- several Λ-terms modifiers (type changes), that can be:

rigid the same modifier applies once for every occurrence. (blocks impossible copredications)

flexible modifier: a different modifier can be used for each occurrence.

30. Rules

$$\frac{A/B : f^{U \to T} \quad B : x^{U}}{A : (fx)^{T}} \; /E$$

$$\begin{array}{c} \ldots [B : x^{U}] \\ \vdots \\ \dfrac{A : t^{T}}{A/B : \lambda x^{U} t} \; /I \end{array}$$

$$\frac{A/B : f^{\Pi\alpha . U[\alpha] \to T} \quad B : x^{U[V]}}{A : (f\{V\}x)^{T}} \; /E^{*}$$

Correspondence: syntactic rule / semantic counterpart. Instantiation and application are combined. The funciton only partly specifies it argument. Thereafter *le* is an instance of this process.

31. Specialised Lexical Semantics

A rather minimal and schematic model.
Base types *region* and *path*

- cognitively motivated (Jackendoff)
- linguistically motivated (some verbs require one).

height is a function from regions to their vertical coordinate.
Functions *source* and *destination*: they convert a path p to its source region and its destination region.
Predicate *middle*(p, r) where p is a path and r a region $\rightarrow$ distinction between Initial, Median and Final verbs as in Asher an Sablayrolles

Spatial variable *here* position and orientation of the spatial reference point not necessarily the narrators's place, implemented as a succession of values.

Motion verbs: relations between one or more entities and a *path* (possibly implicit)

Verbs specify lexically which of their arguments follow this path (subject, object or both, see e.g. Nam).

The lexicon specifies which transformations can take place, allowing us to account for contrasts in grammaticality such as the following.

(22) The road leads us to Pau.

(23) *The road accompanies us to Pau.

32. A lexicon

word/phrase syntactic type	lambda-term
chemin *n*	$\lambda x^{immobile\_object} \begin{array}{\|c\|}\hline \\ \hline chemin(x) \\ \hline\end{array}$
g *n/n*	$\lambda P^{immobile\_object \rightarrow t} \lambda p^{path} \begin{array}{\|l\|}\hline x^{immobile\_object} \;\; q^{path} \;\; here^{region} \\ \hline path\_of(y,p) \\ subpath(q,p) \\ source(q) = here \\ \hline\end{array} \oplus (P\ x)$

chemin *n*	$\lambda x^{immobile\_object}$ $\begin{array}{|c|}\hline \quad \\ \hline \mathit{chemin}(x) \\ \hline\end{array}$

"chemin" is true of entities of type immobile object for which *chemin*(x) holds

g n/n	$\lambda P^{immobile\_object \to t}\, \lambda p^{path} \begin{array}{|l|}\hline x^{immobile\_object} \quad q^{path} \quad here^{region} \\ \hline path\_of(y, p) \\ subpath(q, p) \\ source(q) = here \\ \hline \end{array} \oplus (P\ x)$

coercion g: from an immobile object x to a path p, correspondence indicated by the predicate *path_of*
and selecting a sub-path q of p going forward from *here*, which may or may not go to the end of the path p.
both x (immobile) and p (path) are DRT referents. Indeed, both aspect can be used “a brick road to Pau” and anaphors can refer to either aspect.

(24) The street was completed in 1825 (...)

(25) It runs from the Regent’s residence at Carlton House (...) to All Souls Church.

le $(s/(np\backslash s))/n$	$\Lambda\alpha\lambda P^{\alpha\to t}\lambda Q^{\alpha\to event\to t}\lambda e^{event} \begin{array}{\|c\|}\hline x^{\alpha} \\ \hline \\ \hline \end{array} \oplus (P\ x) \oplus ((Q\ x)\ e)$

"Le" is viewed as a generalised quantifier. It selects a noun (subset of α) and a VP which applies to α to produce a sentence.

descend $np\backslash s$	$\lambda x^{person} \lambda e^{event} \begin{array}{\|l\|}\hline p^{path} \\ \hline \mathit{travel}(e,x,p) \\ \mathit{height}(\mathit{source}(p)) > \mathit{height}(\mathit{destination}(p)) \\ \hline\end{array}$
h $(np\backslash s)/(np\backslash s)$	$\lambda P^{person \to event \to t} \lambda p^{path} \lambda e^{event} \begin{array}{\|l\|}\hline \\ \hline \begin{array}{\|l\|}\hline x^{person} \\ \hline \mathit{travel}(e,x,p) \\ \hline\end{array} \Rightarrow ((P\ x)\ e) \\ \hline\end{array}$

"descend" main term:
given a person argument x and an event argument e,
the DRS checks that there exists a path p such that x follows p and that the height at the start of this path p is greater than his height at the end of it.

What about coercion?

The coercion *h* (for "descend" but not for any motion verb) applied to *descend* yields if a person follows the path p, then he descends.

h descend
$np\backslash s$

$$\lambda p^{path} \lambda e^{event} \begin{array}{|l|} \hline \\ \hline \begin{array}{|l|} \hline x^{person} \\ \hline \mathit{travel}(e, x, p) \\ \hline \end{array} \Rightarrow \begin{array}{|l|} \hline \\ \hline \mathit{height}(\mathit{source}(p)) > \\ \mathit{height}(\mathit{destination}(p)) \\ \hline \end{array} \\ \hline \end{array}$$

Note that "h descend" does not commit us to concluding that anyone actually takes the path. This must be deduced separately.

le{path} (g chemin) $\lambda P^{path \rightarrow event \rightarrow t} \lambda e^{event}$

$s/(np \backslash s)$

$y^{immobile\_object}$ p^{path} q^{path} $here^{region}$
$chemin(y)$ $path\_of(y,p)$ $subpath(q,p)$ $source(q) = here$

" le chemin" with type assignment $np - \iota x^{immobile\_object}.chemin(x)$ does not combine with "descend" which requires a person as its argument, $np \backslash s - \lambda y^{person}$ ...

Both “chemin” and “descend” permit lexically anchored type coercions, which solves the type mismatch:

- “chemin” has a lexical lambda term g which coerces it in such a way that “le chemin” obtains type assignment $np - \iota x^{path}.chemin(x)$
- whereas “descend” has a lexical lambda term h which coerces its lexical semantics to $np \backslash s - \lambda y^{path}$....
- With both coercions “le chemin descend” is a correctly typed term, with “le chemin” being a term of type $path$ and “descend” a term of type $path \rightarrow t$.

33. Two remarks

variable *here* (place + orientation)
no incoherence between “le chemin monte” and “le chemin descend” (reversed orientation)
“pendant deux heures” simple Davidsonian analysis: the duration of the corresponding event is two hours.

pendant 2 h. $\lambda s^{event \rightarrow t} \lambda e^{event} (s\ e) \oplus$

duration$(e, 2h)$

$s \backslash s$

34. Some reflections

- System F is a very simple — but powerful — extension to the simply typed lambda calculus. This speaks in favor of our approach, provided we can account for the same data as Asher (2011).
- Do we need subtyping? A flat ontology seems rather unsatisfactory. However, coherently extending system F with subtyping is not straightforward.
- System F plus subtyping plus restricted quantification over subtypes of a type?

Part V

Towards Segmented Discourse Representation Structures

35. Objective

Finding discourse relations,

- in particular *narration* and *elaboration* for the itinerary
- others, to skip what is irrelevant.

36. Some phenomena

(26) nous descendons, pendant un quart d'heure, la vallée de l'Esera.
we descend, for a quarter of an hour, the Esera valley.

(27) La lune, qui éclaire notre marche, nous fait découvrir sur la droite un sentier qui serpente.
The moon, which lightens our steps, allows us to discover a winding path on our right.

(28) Il nous conduit sur un petit plateau, au milieu de sapins, au-dessus et à quelque distance du torrent de Ramun.
It leads us to a small plateau, surrounded by firs, at some distance of and above the Ramun torrent.

"Il" (it) in sentence 28 refers to "un sentier qui serpente"
imposes anaphora resolution before coercion.
Constraints on the possible interpretations *Background*(26,27) and *Narration*(27,28).

37. Examples, yet other remarks

Rhetorical structure: important but hard to infer.

(29) Nous partimes pour Barèges à 8 heures du matin par une fort jolie route qui nous conduisit à Lourdes.
We left (PS) for Barèges at 8 in the morning, taking a very pretty road which led (PS) us to Lourdes.

(30) (...) qui va en se resserrant jusqu'à Pierrefite, où les routes de Lux et de Cauterets séparent.
(...) which goes shrinking along the way, up to Pierrefite, where the roads to Lux and to Cauterets split.

(31) Celle de Lux entre dans une gorge qui vous mène au fond d'un précipice et traverse le gave de Pau.
The one to Lux enters a gorge which leads you to the bottom of a precipice and traverses the Gave de Pau.

(32) (...) Après une longue marche, l'on arrive à Barèges à 6 heures du soir.
(...) After a long walk, we arrive in Barèges at 6 in the evening.

29 introduces the destination and therefore the whole spatio-temporal extension route. The following will therefore constitute an

Elaboration relation between this sentence and the sequence of 30-32.

It is (at first sight) difficult to decide on the discourse relation of Sentence 31: it would certainly be possible to have a later phrase beginning with “Celle de Cauterets” (the road leading to Cauterets) and a number of the following sentences (omitted here for space reasons) give further background information about the road to Lux.

However, at sentence 32, it suddenly becomes evident that the author has been describing the road while following it.

38. Conclusions and Future Work

Model of "virtual movement" in a type-logical grammar by extending Montague-style semantics

- DRT
- Generative Lexicon

Using system F for varying types and flexibility when it is lexically allowed.
Implementation of coercion in a prototype by Emeric Kien.
Open (common) problems

- anaphora resolution
- determining the appropriate discourse relations between segments of text

Suppressing DRS who are not relevant to the itinerary question (comparison, reason why they acted that way).