



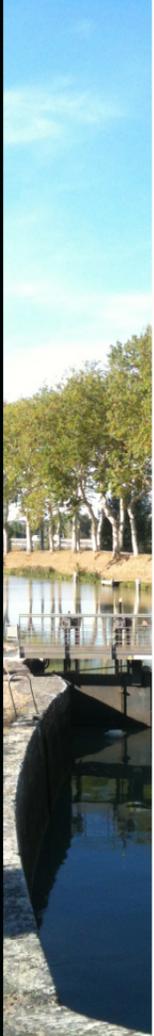
Présentation du Peps HuMain CoLan (LaBRI, LPL, LIRMM, IML)

COMplexité et LANGage
une etude formelle et expérimentale des
mécanismes de compréhension

Christian Retoré

Université de Bordeaux LaBRI

Workshop Slam: (in)cohérence du discours
LORIA, Nancy, 12 & 13 décembre 2013



A Objective of the project



A.1. Who

Laboratoire Parole et Langage (Aix)

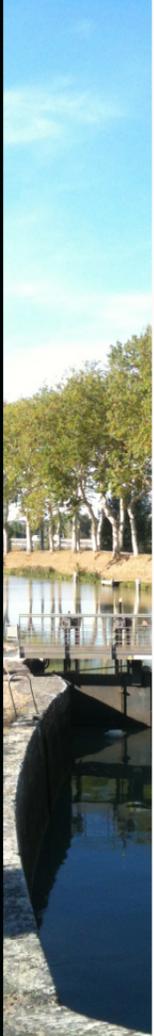
Ph. Blache, M. Champagne-Lavau, S. Dufour, L. Prévot, S. Rauzy

Laboratoire Bordelais de Recherche en Informatique (Bordeaux)

R. Moot, Ch. Retoré (porteur)

LIRMM Montpellier J.-Ph. Prost

IML Marseille M. Quatrini



A.2. Questions under discussions

Human processing, and especially its complexity:

What is difficult to understand and why is it so?

Selected questions in **semantics**:

- coreference (syntax and semantics) (skipped in this talk)
- quantification
- meaning transfers

A.3. Examples: meaning transfers

Selectional restriction meaning transfers, coercions

- (1) # A chair barked.
- (2) Liverpool is a big place.
- (3) Liverpool won the cup.
- (4) Liverpool voted against having a mayor.

Felicitous and infelicitous copredications

- (5) Liverpool is a big place and voted against having a mayor.
- (6) # Liverpool won the cup and voted against having a mayor.



A.4. Quantification

- (7) A secretary will introduce any student to some professor.
- (8) The Brits love France.
- (9) The lion intends to escape.
(a given lion or the generic lion or the set of lions or a set of standard properties of lions or...)
- (10) *In France, every student has an INE number.
does one imagine a student, a set, a student he knows*

A.5. Questions preferably NOT under discussions

Study of the language understanding and not of the reasoning task.

Some works study some phenomena (e.g. quantifiers) by model checking and formulae (not sentences) could be given instead of natural language expression: not relevant to our project.

However language understanding can only be observed through other tasks.



A.6. How to observe difficulty in understanding

Corpus → inappropriate (how do you observe misunderstandings?)

Dialog → why not but phenomena under discussions are rare (corpus CID, "insolite")

Experiments:

- Reaction time (time to click or EEG)
- Eye tracking on sentences and pictures
- Later fMRI

Model checking should simple

(the computation time should be negligible).

As less context and pragmatics as possible.

A.7. Extract complexity measures from formalisation

Some complexity measures relevant to human syntax processing:

LR parsers, stack size of automata, dependency in DG, crossing axioms

In CG (M. Johnson, G. Morrill) rather **lost in parsing** when

one is waiting for several phrases (the superfluous phrases are less problematic)

especially if one waits for several similar phrases (e.g. several verbs).



A.8. Extract complexity measures of semantic processing from formalisation

- Multiple quantifiers, especially with alternation (like center embedded relatives: if more than two, we are lost).
- Order of the logic:
nested implications: if (if A then B) then C
- Order of the quantification process:
quantifying over individuals, sets of individuals, ...
- Generics (set / standard individual) (a student that works succeeds)
- Existential qualifiers are easier than universals.
- Relation between meanings (logical polysemy): derived meanings
- **Ambiguity (several readings) → complexity.**
- What else?



A.9. Semantics / pragmatics

People intend to understand something and to develop possible situation when they hear a sentence even out of context.

This could be called semantics, as opposed to pragmatics which involves discursive context, world knowledge, reasoning etc.

Admittedly, the border is quite thin.



**B All-time classics:
reminder Montague semantics**



B.1. Logic

Logos, logic, philosophy of language, semantics...

Difficult to tell the difference!

From the beginning two related parts:

- nowadays called lexical semantics: interpreting terms (words, noun phrases, even **quantified** nouns phrases)
- nowadays called formal semantics: interpreting propositions, reasoning

B.2. Computing logical forms à la Montague

Mind that there are TWO logics: composition / logical form:

One for expressing meanings:

formulae of first or higher order logic, single or multi sorted.

One for meaning assembly:

proofs in intuitionistic propositional logic, λ -terms expressing the well-formedness of formulae.

Both may introduce some complexity in processing (human processing).

B.3. Representing formulae within lambda calculus — connectives

Assume that the base types are

e (individuals, often there is just one) and

t (propositions)

and that the only constants are

the logical ones (below) and

the relational and functional symbols of the specific logical language (on the next slide).

Logical constants:

- \sim of type $t \rightarrow t$ (negation)
- $\supset, \&, +$ of type $t \rightarrow (t \rightarrow t)$
(implication, conjunction, disjunction)
- two constants \forall and \exists of type $(e \rightarrow t) \rightarrow t$

B.4. Representing formulae within lambda calculus — language constants

The language constants for multi sorted First Order Logic:

- R_q of type $e \rightarrow (e \rightarrow (\dots \rightarrow e \rightarrow t))$
- f_q of type $e \rightarrow (e \rightarrow (\dots \rightarrow e \rightarrow e))$

one two-place predicate	
<i>likes</i>	$\lambda x^e \lambda y^e (\underline{\text{likes}}^{e \rightarrow (e \rightarrow t)} y) x$
two one place predicates	
<i>cat</i>	$\lambda x. \underline{\text{cat}}^{e \rightarrow t}$
<i>sleeps</i>	$\lambda x. \underline{\text{sleep}}^{e \rightarrow t}$
two proper names	
<i>Evora</i>	$\underline{\text{Evora}} : e$
<i>Anne–Sophie</i>	$\underline{\text{Anne–Sophie}} : e$

possibly $(e \rightarrow t) \rightarrow t$

Normal terms (preferably η -long) of type t are formulae.

B.5. Montague semantics. Syntax/semantics.

(Syntactic type)*	=	Semantic type
S^*	=	t a sentence is a proposition
np^*	=	e a noun phrase is an entity
n^*	=	$e \rightarrow t$ a noun is a subset of the set of entities
$(A \setminus B)^* = (B/A)^*$	=	$A \rightarrow B$ extends easily to all syntactic categories of a Categorical Grammar e.g. a Lambek CG

B.6. Montague semantics. Algorithm

1. Replace in the lambda-term issued from the syntax the words by the corresponding term of the lexicon.
2. Reduce the resulting λ -term of type t its normal form corresponds to a formula, the "meaning".



B.7. Ingredients: a parse structure & a lexicon

Syntactical structure

(some (club)) (defeated Leeds)

Semantical lexicon:

word	<i>semantics : λ-term of type (sent. cat.)*</i> <i>x^v the variable or constant x is of type v</i>
<i>some</i>	$(e \rightarrow t) \rightarrow ((e \rightarrow t) \rightarrow t)$ $\lambda P^{e \rightarrow t} \lambda Q^{e \rightarrow t} (\exists^{(e \rightarrow t) \rightarrow t} (\lambda x^e (\wedge^{t \rightarrow (t \rightarrow t)} (P x)(Q x))))$
<i>club</i>	$e \rightarrow t$ $\lambda x^e (\text{club}^{e \rightarrow t} x)$
<i>defeated</i>	$e \rightarrow (e \rightarrow t)$ $\lambda y^e \lambda x^e ((\text{defeated}^{e \rightarrow (e \rightarrow t)} x)y)$
<i>Leeds</i>	e Leeds

B.8. Computing the semantic representation

- 1) Insert the semantics terms into the parse structure
- 2) β reduce the resulting term

$$\begin{aligned} & \left((\lambda P^{e \rightarrow t} \lambda Q^{e \rightarrow t} (\exists^{(e \rightarrow t) \rightarrow t} (\lambda x^e (\wedge (P x) (Q x)))) (\lambda x^e (\text{club}^{e \rightarrow t} x))) \right) \\ & \quad \left((\lambda y^e \lambda x^e ((\text{defeated}^{e \rightarrow (e \rightarrow t)} x) y)) \text{Leeds}^e \right) \\ & \quad \quad \quad \downarrow \beta \\ & (\lambda Q^{e \rightarrow t} (\exists^{(e \rightarrow t) \rightarrow t} (\lambda x^e (\wedge^{t \rightarrow (t \rightarrow t)} (\text{club}^{e \rightarrow t} x) (Q x)))))) \\ & \quad (\lambda x^e ((\text{defeated}^{e \rightarrow (e \rightarrow t)} x) \text{Leeds}^e)) \\ & \quad \quad \quad \downarrow \beta \\ & (\exists^{(e \rightarrow t) \rightarrow t} (\lambda x^e (\wedge (\text{club}^{e \rightarrow t} x) ((\text{defeated}^{e \rightarrow (e \rightarrow t)} x) \text{Leeds}^e)))) \end{aligned}$$

Usually human beings prefer to write it like this:

$$\exists x : e (\text{club}(x) \wedge \text{defeated}(x, \text{Leeds}))$$

B.9. Montague: good architecture / limits

Good trick (Church):

a propositional logic for meaning assembly (proofs/ λ -terms)
computes

formulae of another logic HOL / FOL (formulae/meaning; no
proofs)

reification for remaining in FOL can be discussed

The dictionary says "barks" requires a subject of type "animal".
How could we block:

(11) * The chair barked.

By type mismatch, $(f^{A \rightarrow X}(u^B))$ hence **many types** are needed.

Description with few operators

→ **factorise** similar operations acting on terms/types

→ **quantification over types**



C $\wedge T_{y_n}$: **system F** tuned for semantics

C.1. System F

Types:

- t (prop)
- many entity types e_i
- type variables α, β, \dots
- $\Pi\alpha. T$
- $T_1 \rightarrow T_2$

Terms

- Constants and variables for each type
- $(f^{T \rightarrow U} a^T) : U$
- $(\lambda x^T. u^U) : T \rightarrow U$
- $t^{(\Lambda\alpha. T)}\{U\} : T[U/\alpha]$
- $\Lambda\alpha. u^T : \Pi\alpha. T$ — no free α in a free variable of u .

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).

C.2. Basic facts on system F

Logician / philosopher often ask whether system F is safe?

We do not really need system F but any type system with quantification over types. F is syntactically the simplest. (polynomial Soft Linear Logic of Lafont is enough)

Confluence and strong normalisation — requires the comprehension axiom for all formulae of HA_2 . (Girard 1971)

A concrete categorical interpretation with coherence spaces that shows that there are distinct functions from A to B .

Terms of type t with constants of multisorted FOL (resp. HOL) correspond to multisorted formulae of FOL (resp. HOL)

C.3. Examples of second order usefulness

Arbitrary modifiers: $\Lambda\alpha\lambda x^A y^\alpha f^{\alpha\rightarrow R}.((\text{read}^{A\rightarrow R\rightarrow t} x) (f y))$

Polymorphic conjunction:

Given predicates $P^{\alpha\rightarrow t}$, $Q^{\beta\rightarrow t}$ over respective types α , β ,
given any type ξ with two morphisms from ξ to α and to β
we can coordinate the properties P , Q
of (the two images of) an entity of type ξ :

The polymorphic conjunction $\&^\Pi$ is defined as the term

$$\begin{aligned} \&^\Pi = \Lambda\alpha\Lambda\beta\lambda P^{\alpha\rightarrow t}\lambda Q^{\beta\rightarrow t} \\ \Lambda\xi\lambda x^\xi\lambda f^{\xi\rightarrow\alpha}\lambda g^{\xi\rightarrow\beta}. \\ (\text{and}^{t\rightarrow t\rightarrow t} (P (f x))(Q (g x))) \end{aligned}$$

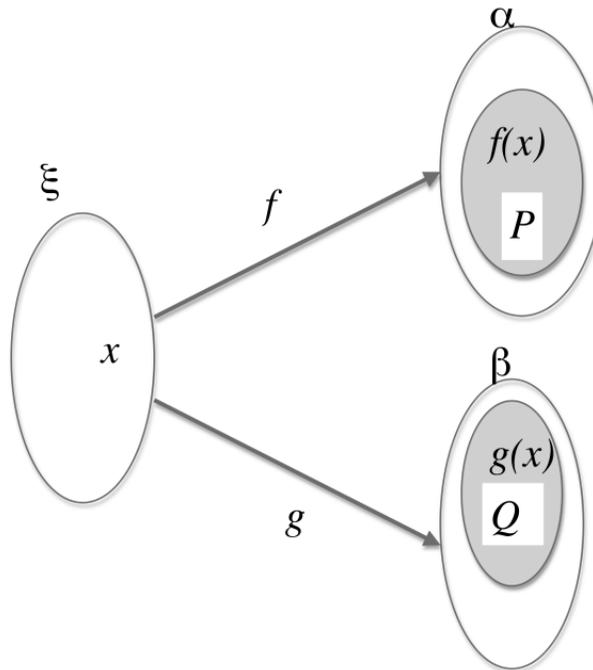


Figure 1: Polymorphic conjunction: $P(f(x)) \& Q(g(x))$
with $x : \xi$, $f : \xi \rightarrow \alpha$, $g : \xi \rightarrow \beta$.

C.4. Coercive subtyping for F (Luo, Soloviev, Retoré)

Ontological inclusions: *A car is a vehicle*

Key property: at most one coercion between any two types.

Given coercions between base types.

Propagates through type hierarchy (unique possible restoration).

$$\text{coercive application} \quad \frac{f : A \rightarrow B \quad u : A_0 \quad A_0 < A}{(f \ a) : B}$$

$$\begin{array}{c} \dots\dots\dots \\ \frac{A < B \quad C < D}{B \rightarrow A < C \rightarrow D} \quad \frac{A < B}{X \rightarrow A < X \rightarrow B} \quad \frac{A < B}{B \rightarrow X < A \rightarrow X} \\ \dots\dots\dots \\ \frac{S[X] < T[X]}{\Pi X. S[X] < \Pi X. T[X]} \quad \frac{U < T[X]}{U < \Pi X. T[X]} \text{ no free } X \text{ in } U \quad \frac{S[W] < U}{\Pi X. S[X] < U} \\ \dots\dots\dots \\ \frac{U < \Pi X. T[X]}{U < T[A]} \quad \frac{\Pi X. S[X] < U}{S[A] < U} \end{array}$$

Key lemma: transitivity of $<$ is unnecessary.



D System F based semantics and pragmatics



D.1. Examples

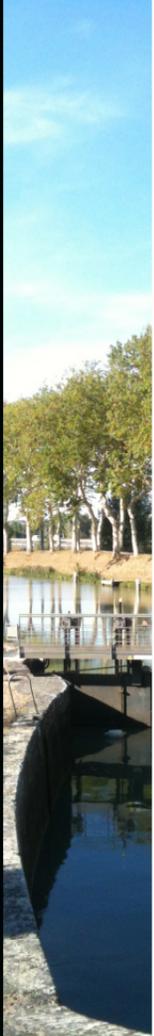
- (12) Dinner was delicious but took ages.
(event / food)
- (13) * The salmon we had for lunch was lightning fast.
(animal / food)
- (14) I carried the books from the shelf to the attic.
Indeed, I already read them all.
(phys. / info — think of possible multiple copies of a book)
- (15) Liverpool is a big place and voted last Sunday.
(geographic / people)
- (16) * Liverpool is a big place and won last Sunday.
(geographic / football club)

D.2. Counter-examples

Nevertheless:

- (17) Barcelona won four champions leagues and organised the olympiads.
(Delofeu)
- (18) Libourne, a small town nearby Bordeaux, defeated Lille.

Contextualisation, contrast,... out of reach for the time being!



D.3. Types and terms: system F

System F with many base types e_i (many sorts of entities)

v (for events who play a particular role)

t truth values

types variables roman upper case, greek lower case

usual terms that we saw, with constants (free variables that cannot be abstracted)

Every normal terms of type t with free variables being logical variables (of a the corresponding multi sorted logic L) correspond to a formula of L .

D.4. The Terms: principal or optional

A standard λ -term attached to the main sense:

- Used for compositional purposes
- Comprising detailed typing information (restrictions of selection)

Some optional λ -terms (none is possible)

- Used, or not, for adaptation purposes
- Each associated with a constraint : *rigid*, \emptyset

Both function and argument may contribute to meaning transfers.

Weighted morphisms can be a measure of complexity.





D.5. RIGID vs FLEXIBLE use of optional terms

RIGID

Such a transformation is exclusive:

the other aspects of the same word are not used.

Each time we refer to the word it is with the same aspect.

FLEXIBLE

There is no constraint.

Any subset of the flexible transformation can be used:

different aspects of the words can be simultaneously used.

D.6. Correct copredication

word	principal λ -term	optional λ -terms	rigid/flexible
<i>Liverpool</i>	<i>liverpool</i> ^T	$Id_T : T \rightarrow T$ (F) $t_1 : T \rightarrow F$ (R) $t_2 : T \rightarrow P$ (F) $t_3 : T \rightarrow Pl$ (F)	
<i>is_a_big_place</i>	<i>big_place</i> : $Pl \rightarrow \mathbf{t}$		
<i>voted</i>	<i>voted</i> : $P \rightarrow \mathbf{t}$		
<i>won</i>	<i>won</i> : $F \rightarrow \mathbf{t}$		

where the base types are defined as follows:

- T Town
- F football club
- P people
- Pl place

D.7. Meaning transfers

- (19) Liverpool is a big place.
- (20) Liverpool won.
- (21) Liverpool voted.

$big_place^{Place \rightarrow t} Liverpool^{Town}$

Type mismatch, use the appropriate optional term.

$big_place^{Place \rightarrow t} (t_3^{Town \rightarrow Place} Liverpool^{Town})$

D.8. (In)felicitous copredications

Use polymorphic "and"... specialised to the appropriate types:

- (22) Liverpool is a big place and voted.

Town → *Place* and *Town* → *People*

fine

- (23) * Liverpool won and voted.

Town → *FootballClub* and *Town* → *People*

blocked because the first transformation is **rigid**.

(sole interpretation: *football* team or committee voted)

D.9. Liverpool is a big place

Type mismatch:

$$big_place^{Pl \rightarrow t}(Liverpool^T)$$

big_place applies to “places” (type Pl) and not to “towns” (T)

Lexicon $t_3^{T \rightarrow Pl}$ turns a town (T) into a place (Pl)

$$big_place^{Pl \rightarrow t}(t_3^{T \rightarrow Pl} Liverpool^T)$$

only one optional term, the (F)/ (R) difference is useless.

D.10. Liverpool is a big place and voted

Polymorphic AND yields: $(\&^\Pi(\text{big_place})^{Pl \rightarrow \mathbf{t}}(\text{voted})^{P \rightarrow \mathbf{t}})$

Forces $\alpha := Pl$ and $\beta := P$, the properly typed term is

$$\&^\Pi\{Pl\}\{P\}(\text{is_wide})^{Pl \rightarrow \mathbf{t}}(\text{voted})^{P \rightarrow \mathbf{t}}$$

It reduces to:

$$\Lambda \xi \lambda x^\xi \lambda f^{\xi \rightarrow \alpha} \lambda g^{\xi \rightarrow \beta} (\text{and}^{\mathbf{t} \rightarrow \mathbf{t}})^{\rightarrow \mathbf{t}} (\text{is_wide } (f \ x)) (\text{voted } (g \ x)))$$

Syntax applies it to “*Liverpool*” so $\xi := T$ yielding

$$\lambda f^{T \rightarrow Pl} \lambda g^{T \rightarrow P} (\text{and } (\text{is_wide } (f \ \text{Liverpool}^T)) (\text{voted } (g \ \text{Liverpool}^T))).$$

The two flexible optional λ -terms $t_2 : T \rightarrow P$ and $t_3 : T \rightarrow Pl$ yield

$$(\text{and } (\text{is_wide}^{Pl \rightarrow \mathbf{t}} (t_3^{T \rightarrow Pl} \ \text{Liverpool}^T)) (\text{voted}^{Pl \rightarrow \mathbf{t}} (t_2^{T \rightarrow P} \ \text{Liverpool}^T)))$$

D.11. Liverpool voted and won

As previously but with *won* instead of *big_place*.

The term is:

$$\lambda f^{T \rightarrow P} \lambda g^{T \rightarrow P} (\text{and } (won (f \text{ Liverpool}^T)) (voted (g \text{ Liverpool}^T))))$$

for “*won*”, we need to use the transformation $t_1 : T \rightarrow F$

but T_1 is rigid, hence we cannot access to the other needed transformation into a “*place*”.



E Quantification and other questions related to lexical semantics

E.1. Quantifier: critics of the standard solution

Syntactical structure of the sentence \neq logical form.

(24) Keith played some Beatles songs.

(25) syntax (Keith (played (some (Beatles songs))))

(26) semantics: (some (Beatles songs)) (λx . Keith played x)

Asymmetry class / predicate

(27) Some politicians are crooks

(28) ? Some crooks are politicians

(29) $\exists x$. *crook*(x) & *politician*(x)

There can be a reference before the predicate arrives (if any):

(30) Un luth, une mandore, une viole, que Michel-Ange... (M. Énard)

E.2. A solution: Hilbert's epsilon

$\varepsilon : \Lambda\alpha(\alpha \rightarrow \mathbf{t}) \rightarrow \alpha$ with $F(\varepsilon_x F) \equiv \exists x. F(x)$.

A cat. $cat^{animal \rightarrow \mathbf{t}} \quad (\varepsilon\{animal\}cat^{animal \rightarrow \mathbf{t}}) : animal$

Presupposition $F(\varepsilon_x F)$ is added: $cat(\varepsilon\{animal\}cat^{animal \rightarrow \mathbf{t}})$

$\varepsilon_x F$: individual. Follows syntactical structure. Asymmetry subject/predicate.

ε also solves the so-called E-type pronouns interpretation:

(31) A man came in. He sat down.

(32) "He" = "A man" = $(\varepsilon_x M(x))$.

For applying ε to a type say cat ,

any type has a predicative counterpart cat (type) $\widehat{cat} : \mathbf{e} \rightarrow \mathbf{t}$.

(domains can be restrained / extended)

E.3. Remarks on ε

Hilbert's work: fine! (Grundlagen der Mathematik, with P. Bernays)

Rule 1: From $P(x)$ with x generic infer $P(\varepsilon_x.\neg P(x)) \equiv P(\tau_x.P(x)) \equiv \forall x P(x)$

Rule 2: From $P(t)$ infer $P(\varepsilon_x P(x)) \equiv \exists x P(x)$

ε -elimination (1st & 2nd ε -theorems), proof of Herbrand theorem.

Little else is known (extra formulae, proofs, models), erroneous results.

$Sleeps(\varepsilon_x Cat(x)) \equiv ???$

$(Cat(\varepsilon_x Cat(x)) \& Sleeps(\varepsilon_x Cat(x))) \equiv \exists x Cat(x) \& Sleeps(x)$

Heavy notation: $\forall x \exists y P(x, y)$ is $P(\tau_x P(x, \varepsilon_y P(\tau_x P(x, y), y)), \varepsilon_y P(\tau_x P(x, y), y))$

von Heusinger interpretations differ for different occurrences of $\varepsilon_x F(x)$.

- (33) a. A tall man went in. A blonde man went out.
b. *Not the same F but necessarily different interpretations.*

E.4. Intuitive interpretation and logic: some perspectives

Cohabitation of types and formulae of first/higher order logic:

Typing (\sim presupposition) is irrefutable $sleeps(x : cat)$

Type to Formula:

type *cat* mirrored as a predicate $\widehat{cat} : e \rightarrow t$

Formula to Type?

Formula with a single free variable \sim type?

$cat(x) \wedge belong(x, john) \wedge sleeps(x) \sim$ type?

At least it is not a natural class.

E.5. Other phenomena handled by the same model

Generalised quantifiers (“*most*”) with generic elements.
The Brits love France.

Plurals: collective / distributive readings (with Moot)
(The players from) Benfica won although they had the flu.

Virtual traveller / fictive motion (with Moot & Prévot)
“The road does down for twenty minutes”

Deverbals: meanings copredications (with Livy Real):
*“A assinatura atrasou três dias / * e estava ilegível.”*



**F Wide (wild?) semantic model vs.
possible experiments**

F.1. Experiments deserve a lot of care

A single phenomena at a time.

As little pragmatics as possible.

Respect the protocol e.g. understanding of quantifiers differ between 18 and 25.

We chose a very simple experiment to start with:

difference of scope taking between "chaque" and "tous les".

(34) Un rond est relié à tous les carrés.

(35) Un rond est relié à chaque carré.

(36) Un rond est relié à tous les coins du carré.

(37) Un rond est relié à chaque coin du carré.

The choice of the predicate has some effect.

The judgement differ w.r.t. to the individuals.

F.2. Lexical semantics pragmatics

Idea the more the meaning transfer is

unusual

far from the initial sense

(the two can go together) the more difficult is the interpretation using it.

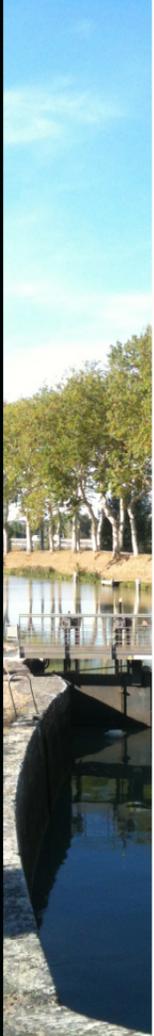
- (38) I finish my book
("read" < "write" << "bind")
- (39) My car has been written out.

F.3. Link with Slam

Lexical similarity and disruption is a well known way in Natural Language Processing to structure a discourse, by finding the cohesive parts and the disruptions. They can also be endowed with some psychological interpretations.

Also the absence of perception of some ambiguities or the perception of may deserve some psychological interpretation.





G Conclusion



G.1. Filling the gap

After a few months we find a big gap between what we initially wanted to check with experiments and the experiments that can be done and properly interpreted.

One can do much more complex experiments but then the difficulty is to interpret the results.

So we ought to be patient.



Any question?