

# **Compositionality and the Lexicon**

Bruno Mery

2009, April 10th

*Lexique organisé pour la composition sémantique*

LaBRI, équipe Méthodes Formelles  
Projet Signes – Université de Bordeaux

# Outline

- State of the Lexicon
- Types and Terms
- Formulæ and Models

## Lexical Semantics

- A not-so-recent problem: single-sense lexicons
- Sense disambiguation and lexical semantics
- Linguistic and background knowledge
- The advent of the Generative Lexicon
- A gap within the formalism

## Classical GL examples

- Qualia
  - *A quick cigarette* (telic)
  - *A partisan article* (agentive)
- Dot Objects
  - *An interesting book* (*I*)
  - *An heavy book* ( $\varphi$ )
  - *A large city* (*T*)
  - *A cosmopolitan city* (*P*)
- Co-predications
  - *A heavy, yet interesting book*
  - *Paris is a large, cosmopolitan city*
  - ? *A fast, delicious salmon*
  - ?? *Washington is a small city and signed a trade agreement with Paris*

## Our principles

- Remain within reach of Montagovian compositional semantics
- Allow both predicate and argument to contribute lexical information to the compound
- Integrate within existing discourse models

In [BMRBP], we advocate a system based on *optional modifiers*.

## Overview of the Lexicon

How much information should a lexicon store ?

- Basic compositional data: number, type, optional character of arguments
- Lexical data for adaptations: qualia, dot objects...
- Constraints on modifiers induced by lexical data
- Interpretation(s) of each term

# The Type Language

- Montagovian composition:
    - Each predicate indicates the typing and expected order for its arguments
  - GL-style concept hierarchy:
    - Types are different for every distinct lexical behavior
    - A structure (typically, an ontology) details the specialization relations between types
    - The result is close to a language-independent hierarchy of concepts
- Second-order typing, à la [GTL89] is needed for the integration of arbitrary modifiers:*
- $$\Lambda\alpha\lambda x^A y^\alpha f^{\alpha \rightarrow R}. ((\text{read}^{A \rightarrow R \rightarrow t} x) (f y))$$

# The Term Language

Each term comprises:

- A standard  $\lambda$ -term attached to the main sense:
  - Used for compositional purposes
  - Comprising detailed typing information
  - Including slots for optional modifiers
  - $\Lambda\alpha\beta\lambda x^\alpha y^\beta f^{\alpha \rightarrow A} g^{\beta \rightarrow F}.((\text{eat}^{A \rightarrow F \rightarrow t} (f x)) (g y))$
  - Paris<sup>T</sup>
- A set of optional morphisms:
  - Each a one-place predicate
  - Used, or not, for adaptation purposes
  - Each associated with a constraint : *local*, *global*,  $\emptyset$
  - $\left( \frac{Id^{F \rightarrow F}}{\emptyset}, \frac{f_{grind}^{Living \rightarrow F}}{global} \right)$
  - $\left( \frac{Id^{T \rightarrow T}}{\emptyset}, \frac{f_L^{T \rightarrow L}}{\emptyset}, \frac{f_P^{T \rightarrow P}}{\emptyset}, \frac{f_G^{T \rightarrow G}}{global} \right)$

Plus interpretations.

## A Lexical Entry

Every lexeme is associated to an  $n$ -uple such as:

$$\left( \text{Paris}^T, \frac{\lambda x^T. x^T}{\emptyset}, \frac{\lambda x^T. (f_L^{T \rightarrow L} x)}{\emptyset}, \frac{\lambda x^T. (f_P^{T \rightarrow P} x)}{\emptyset}, \frac{\lambda x^T. (f_G^{T \rightarrow G} x)}{global} \right)$$

## Importing an existing GL

- Main type and argument structure: main  $\lambda$ -term
- *Qualia*-roles: local modifiers
- Dot objects: local modifiers
- Some specific constructions are global modifiers (e.g. grinding).
- Inheritance structure: local modifier → parent

## The calculus, summarized

- First-order  $\lambda$ -bindings: usual composition
- Open slots: generate all combinations of modifiers available
- As many interpretations as well-typed combinations

*Paris is an populous city by the Seine river*

$$((\Lambda\xi . \lambda x^\xi f^{\xi \rightarrow P} g^{\xi \rightarrow L} . (\text{and} (\text{populous}^{P \rightarrow t} (f x)) (\text{riverside}^{L \rightarrow t} (g x))))$$

$$\{T\} \text{Paris}^T \lambda x^T (f_P^{T \rightarrow P} x) \lambda x^T . (f_L^{T \rightarrow L} x))$$

## Logical Formulae

- Many possible results
- Our choice: classical, higher-order predicate logic
- No modalities ( $\rightarrow$  interpretation)

and(populous( $f_P(\text{Paris})$ ), riverside( $f_L(\text{Paris})$ )))

## Models and Interpretation

- Modularity: no real need to choose
- Multiple formulæ, multiple interpretations
- Locality principle
- A first implementation target: Herbrand Universes
- A straightforward extension to discourse and dialogue

## Reasons of the approach

- Remaining in a practical, compatible framework
- Finding a simple model to multiple sense adaptation, extension, transfer... situations
- Streamlining the cases described by GL
- Keeping modularity and implementation in mind

## Conclusion

- A framework for all kind of lexical operations
- A simple, streamlined model
- Designed for identification of sense adaptations
- Next :
  - Formal inquiries
  - Prototypes and testing

# References

- [BMRBP] Christian Bassac, Bruno Mery, and Christian Retoré. Towards a Type-Theoretical Account of Lexical Semantics. *Journal of Language, Logic, and Information*, TBP. To appear.
- [Gir71] J. Y. Girard. Une extension de l'interprétation de Gödel l'analyse et son application : l'élimination des coupures dans l'analyse et la théorie des types. In Fenstad, editor, *Proceedings Second Scandinavian Logic Symp.*, pages 63–92, North Holland, Amsterdam, 1971.
- [Gir72] J. Y. Girard. Interprétation fonctionnelle et élimination des coupures de l'arithmétique d'ordre supérieur. Thèse de Doctorat d'État, Université Paris VII, 1972.
- [GTL89] Jean-Yves Girard, Paul Taylor, and Yves Lafont. *Proofs and types*. Cambridge University Press, New York, NY, USA, 1989.
- [Jac01] Evelyne Jacquay. *Ambiguïtés lexicales et traitement automatique des langues : modélisation de la polysémie logique et application aux déverbaux d'action ambigus en français*. PhD thesis, Université de Nancy 2, 2001.
- [MBR07a] Bruno Mery, Christian Bassac, and Christian Retoré. A montagovian generative lexicon. In *Formal Grammar*, 2007.
- [MBR07b] Bruno Mery, Christian Bassac, and Christian Retoré. A montague-based model of generative lexical semantics. In Reinhard Muskens, editor, *New Directions in Type Theoretic Grammars*. ESSLLI, Foundation of Logic, Language and Information, 2007.
- [Nun93] Geoffrey Nunberg. Transfers of meaning. In *Proceedings of the 31st annual meeting on Association for Computational Linguistics*, pages 191–192, Morristown, NJ, USA, 1993. Association for Computational Linguistics.