



Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : <http://oatao.univ-toulouse.fr/>
Eprints ID : 12713

The contribution was presented at TALN 2013 :
<http://www.taln2013.org/>

To cite this version : Retoré, Christian *Sémantique des déterminants dans un cadre richement typé*. (2013) In: 20ème conférence du Traitement Automatique des Langues Naturelles (TALN 2013), 17 June 2013 - 21 June 2013 (Les sables d'olonne, France).

Any correspondance concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr

Sémantique des déterminants dans un cadre richement typé

Christian Retoré

IRIT (CNRS, Toulouse) & Université de Bordeaux (LaBRI)

Christian.Retore@irit.fr

RÉSUMÉ

La variation du sens des mots en contexte nous a conduit à enrichir le système de types utilisés dans notre analyse syntaxico-sémantique du français basé sur les grammaires catégorielles et la sémantique de Montague (ou la lambda-DRT). L'avantage majeur d'une telle sémantique profonde est de représenter le sens par des formules logiques aisément exploitables, par exemple par un moteur d'inférence. Déterminants et quantificateurs jouent un rôle fondamental dans la construction de ces formules, et il nous a fallu leur trouver des termes sémantiques adaptés à ce nouveau cadre. Nous proposons une solution inspirée des opérateurs epsilon et tau de Hilbert, éléments génériques qui s'apparentent à des fonctions de choix. Cette modélisation unifie le traitement des différents types de déterminants et de quantificateurs et autorise le liage dynamique des pronoms. Surtout, cette description calculable des déterminants s'intègre parfaitement à l'analyseur à large échelle du français Grail, tant en théorie qu'en pratique.

ABSTRACT

On the semantics of determiners in a rich type-theoretical framework

The variation of word meaning according to the context led us to enrich the type system of our syntactical and semantic analyser of French based on categorial grammars and Montague semantics (or lambda-DRT). The main advantage of a deep semantic analyse is too represent meaning by logical formulae that can be easily used e.g. for inferences. Determiners and quantifiers play a fundamental role in the construction of those formulae and we needed to provide them with semantic terms adapted to this new framework. We propose a solution inspired by the tau and epsilon operators of Hilbert, generic elements that resemble choice functions. This approach unifies the treatment of the different determiners and quantifiers and allows a dynamic binding of pronouns. Above all, this fully computational view of determiners fits in well within the wide coverage parser Grail, both from a theoretical and a practical viewpoint.

MOTS-CLÉS : Analyse sémantique automatique, Sémantique formelle, Compositionnalité.

KEYWORDS: Automated semantic analysis, Formal Semantics, Compositional Semantics.

1 Présentation

Dans le cadre du traitement automatique des langues, on entend plus souvent parler de sémantique distributionnelle, de vecteurs de mots et de fréquences que de sémantique formelle ou compositionnelle. Certes, les approches quantitatives sont plus aisées à mettre en oeuvre et fournissent des outils efficaces mais elles ne répondent pas aux mêmes questions. Les approches quantitatives sont fort utiles en recherche d'information et en classification car elles permettent de dire *de quoi parle* une phrase, une page web, un texte. En revanche elles ne disent pas ce qu'affirme le texte analysé, *qui fait quoi*. Une phrase peut très bien nier quelque chose, et ainsi causer une erreur à un système de recherche d'information (cf. exemple 1). Il faut aussi savoir reconnaître les pronoms, pour répondre à des questions comme *Geach était-il l'élève de Wittgenstein ?* à partir du web où on ne trouve que l'exemple (2).¹

- (1) Mais vérification faite, ce n'était PAS un OURAGAN qui était passé par là.
- (2) Bien qu'IL n'ait JAMAIS suivi l'enseignement académique de CE DERNIER, cependant IL EN éprouva fortement l'influence.

Ainsi, l'analyse sémantique complète et profonde d'une ou plusieurs phrases reste une tâche pertinente dans le traitement automatique des langues. Ce processus est utilement complété par des techniques statistiques, par exemple pour trouver les paragraphes pertinents ou pour définir des préférences contextuelles lorsque plusieurs sens sont possibles. Une fois les formules logiques construites à partir du texte, un moteur d'inférence est capable de dire si une formule ou un énoncé découle de ce qui a été analysé — plus facilement qu'avec des graphes sémantiques.

Les déterminants (voir par exemple (Corblin *et al.*, 2004)) sont un ingrédient important de l'analyse logique d'une phrase. Les déterminants indéfinis correspondent généralement à une quantification existentielle ou, ce qui s'en approche, à l'introduction d'un référent de discours — il peuvent aussi exprimer une propriété notamment en position d'attribut, mais nous ne parlerons pas de ce dernier sens. Les déterminants définis correspondent plutôt à la désignation d'un élément saillant du contexte, mais il arrive qu'ils introduisent un élément de discours. Assez souvent, ils expriment la quantification vague comme "*beaucoup, la plupart*". La quantification universelle est assez rare, et elle peut s'exprimer par un usage générique de "*un, le, la, les*".

D'un point de vue pratique, ce travail se situe dans le cadre des grammaires catégorielles qui sont une approche de la syntaxe très orientée vers la sémantique compositionnelle. En effet, il est assez aisé de déduire de la structure syntaxique proposée par une grammaire catégorielle une représentation du sens sous forme logique. D'ailleurs, à notre connaissance, les deux seuls systèmes produisant une analyse sémantique complète comme une formule logique (une DRS, en fait) sont basés sur les grammaires catégorielles : *Boxer* de Bos (2008) analyse de l'anglais par des *Categorical Combinatory Grammars* tandis que nous utilisons *Grail* de Moot (2010a,b) basé sur les *Multimodal Categorical Grammars*. Dans un cas comme dans l'autre, la grammaire est acquise automatiquement sur corpus annoté. L'acquisition automatique de la grammaire produit un grand nombre de catégories par mot, et un minimum de traitement probabiliste est nécessaire pour ne considérer que les assignations les plus probables lors de l'analyse. Du point de vue sémantique, ces systèmes utilisent la correspondance entre syntaxe et sémantique telle qu'initée

1. Sauf mention contraire, nos exemples proviennent d'Internet.

Mot	type sémantique u^* terme sémantique : λ -terme de type u^* x^v la variable ou la constante x est de type v
<i>un</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>a_battu</i>	$e \rightarrow (e \rightarrow t)$ $\lambda y^e \lambda x^e ((\text{a_battu}^{e \rightarrow (e \rightarrow t)} x)y)$
<i>Leeds</i>	e Leeds

FIGURE 1 – Un lexique sémantique élémentaire

par Montague c.f. (Moot et Retoré, 2012, chapitre 3). Rappelons la brièvement sur un exemple, car c'est le point de départ de nos travaux.

Supposons que l'analyse syntaxique de "*un club a battu Leeds.*" produise "*(un (club)) (a battu Leeds)*" expression dans laquelle la fonction est systématiquement écrite à gauche. Si les termes sémantiques sont ceux du lexique de la figure 1, alors en remplaçant les mots par les termes sémantiques associés on obtient un grand λ -terme, que l'on peut réduire :

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

Ce λ -terme de type t peut être appelé la *forme logique de la phrase* ; il est plus agréable sous un format standard : $\exists x : e (\text{club}(x) \wedge \text{a_battu}(x, \text{Leeds}))$. Nous verrons ci-après que ce traitement standard de la quantification pose problème, notamment avec la sémantique lexicale.

On observera qu'il y a deux logiques à l'oeuvre. La première est le calcul propositionnel intuitionniste dont on n'utilise que les preuves ou λ -termes : elle assemble des formules partielles. La seconde est une logique dont on n'utilise que les formules. Le λ -terme de type t obtenu *in fine* est effectivement une formulation logique du sens en λ -calcul.

2 Déterminants et quantificateurs

Sans surprise, les déterminants considérés sont de deux sortes, définis et indéfinis — nous essaierons d'éviter les pluriels : ils posent d'autres problèmes abordés dans ce cadre par Moot

et Retoré (2011). Logiquement, les déterminants indéfinis s'apparentent à une quantification existentielle dite généralisée lorsqu'il agissent sur une classe ou lorsqu'ils introduisent un nouveau référent de discours — nous n'aborderons pas le cas où ils introduisent une propriété, par exemple dans un groupe nominal attribut. Précisons de suite qu'un travail de formalisation et d'automatisation comme le nôtre ne peut prétendre atteindre la finesse de travaux plus descriptifs comme ceux de Corblin *et al.* (2004), et que nous sommes donc contraints de schématiser, voire d'ignorer, certaines constructions. Considérons quelques exemples :

- (3) a. J'ai senti **un** animal ME TOUCHER LE PIED.
 b. **Un** parent d'élève de maternelle VIENT CHERCHER SON ENFANT en état d'ébriété,² l'enseignant commet-il une faute en remettant l'enfant à ce parent ?
- (4) a. Aujourd'hui, je me suis réveillé en sursaut parce que j'ai senti **quelque chose** ME TOUCHER LE PIED. Il s'est avéré que c'était mon autre pied.
 b. Précisez si **quelqu'un** VIENT CHERCHER L'ENFANT.
- (5) a. Il y avait **une panthère sortie de la cage**. Elle était attachée. **L'animal** a sauté sur moi.
 b. **Un homme** avait menacé la principale du collège de Monts où son fils était scolarisé. **Le parent d'élève** a été condamné hier.
- (6) A la SPA si ont désire adopter **un animal** il faut donner 500€, et on a 24 Heures pour réfléchir si l'on désire **l'animal** ou non.

Les deux premiers exemples (3) sont à mettre en parallèle avec les deux suivants (4) qui correspondent eux-aussi à une quantification existentielle. Dans cette deuxième version, il n'y a que le prédicat principal, que nous avons choisi pour être le même, et il n'y a plus de restriction à une classe d'objets par un nom commun, avec ou sans compléments (\bar{N} syntaxiquement ou $e \rightarrow t$ sémantiquement). Observons que le traitement usuel de la quantification dans une logique non typée à la Frege ne fait aucune distinction entre le nom quantifié et le prédicat principal.

Les déterminants définis ont un rapport avec les déterminants indéfinis, qui souvent les introduisent, comme le montre les exemples (5) Les expressions se correspondent, et idéalement on aimerait que "le X" soit précédé de "un X", comme dans l'exemple (6). En fait, c'est plutôt rare, et les exemples en corpus sont plutôt comme (5) : l'antécédent de l'anaphore associative n'est pas celui qu'on espérerait pour un traitement automatique, quelques inférences sont nécessaires.

2.1 Traitement usuel et critique

L'analyse traditionnelle attribue à l'article indéfini un terme sémantique exprimant une quantification existentielle.

$$(7) \text{ un} : \lambda P^{e \rightarrow t} \lambda Q^{e \rightarrow t} (\exists \lambda x.^t. \&(P x)(Q x)) : (e \rightarrow t) \rightarrow (e \rightarrow t) \rightarrow t$$

$$(8) \text{ quelque chose} : \exists : (e \rightarrow t) \rightarrow t$$

2. Il s'agit sans doute du parent :-)

Les articles définis sont traités différemment : les groupes nominaux qu'ils introduisent sont plutôt vus comme des anaphores, dites associatives, dont on cherche les référents.³ Cette modélisation classique en sémantique formelle ou dans les grammaires catégorielles pose divers problèmes.

Syntaxe et sémantique Les déterminants traités comme des quantificateurs généralisés mettent à mal la correspondance entre syntaxe et sémantique. La structure sémantique (9c) et la structure syntaxique (9b) ne coïncident pas. Les grammaires catégorielles obtiennent une structure syntaxique (9c), laquelle n'a rien de naturel, au prix d'une catégorie syntaxique différente pour chaque position syntaxique du groupe nominal quantifié : cela n'est guère satisfaisant.

- (9) a. elle écoutait une chanson de lassana hawa
b. SYNT. USUELLE : (elle (écoutait (**une** (chanson (de lassana hawa))))
c. SEM. & CG : ((**une** (chanson (de lassana hawa))) (λx elle écoutait x))

Référence du groupe nominal quantifié Comme le fait remarquer (Geach, 1962), on peut se forger une interprétation du groupe nominal défini ou indéfini, avant même que le prédicat principal soit énoncé... et dans l'exemple (10c) il n'arrive jamais.

- (10) a. Ensuite, LES ÉLÈVES sont allés en salle info, pour réaliser un caryotype classé.
b. Ensuite, DES ÉLÈVES sont venus voir ce que l'on faisait.
c. Un luth, une mandore, une viole, que Michel-Ange [...]. [phrase nominale].⁴

Asymétrie entre thème et rhème L'approche standard impose une symétrie entre le prédicat principal et la restriction à une classe d'objets, symétrie que la langue ne fait pas.

- (11) a. Certains politiciens sont des menteurs car ce qui les intéresse (...)
b. * Certains menteurs sont des politiciens car ce qui les intéresse (...)⁵

Définis et indéfinis Comme remarqué dans (Egli et von Heusinger, 1995; von Heusinger, 1997, 2004), l'unicité est loin d'être requise lorsque l'on utilise un déterminant défini. Un locuteur peut dire "l'île" du lac de Constance, alors qu'il y en a trois, comme ici sur Internet :

- (12) Recueilli (...) par les moines de l'abbaye de Reichenau, sur l'île du lac de Constance,

De plus, "un" et "le" se rapprochent aussi car le contexte extra linguistique permet parfaitement d'utiliser l'article défini sans que le référent n'ait jamais été introduit.

- (13) J'avais pris l'assurance 'automatiquement' avec le prêt immobilier lors de l'achat de *la maison*.⁶

Les déterminants "un" et "le" se ressemblent, alors que la sémantique formelle usuelle les oppose. Selon von Heusinger il s'agit d'une différence d'interprétation et non de forme logique : "un" choisit un nouvel élément, tandis que "le" choisit le plus saillant des référents possibles.

3. Une autre approche utilise une fonction de choix, mais nous allons justement présenter une solution de cet ordre.

4. Mathias Enard, *Parle-leur de batailles, de rois et d'éléphants* Actes-Sud, 2010.

5. Cet exemple est de nous, pour faire contraste avec le précédent.

6. Dans ce récit trouvé sur une FAQ il n'a jamais été question de "maison" auparavant

Les pronoms de type E ne sont pas des pronoms particuliers, mais une interprétation possible et très naturelle des pronoms due à Evans (1977). Cette interprétation consiste à associer au pronom le terme sémantique de son antécédent. On peut aussi traiter de la sorte les groupes nominaux introduits par l'article défini (Egli et von Heusinger, 1995; von Heusinger, 1997, 2004). Cela permet d'étendre la portée du quantificateur existentiel souvent introduit par "un" comme le font la DRT et la *dynamic predicate logic*. Ce type d'interprétation n'est pas possible avec le terme sémantique standard associé à "un" en (7).

- (14) a. Soudain, **un homme** est entré.
 b. IL / CET HOMME / L'HOMME a hurlé « Donne-moi la caisse ! ».

3 Opérateurs de Hilbert, quantificateurs et déterminants

Les opérateurs de Hilbert, surtout ι et ϵ , ont été utilisés pour modéliser les déterminants et la quantification existentielle, en particulier par von Heusinger (Egli et von Heusinger, 1995; von Heusinger, 1997, 2004). Ces opérateurs, bien décrits par Hilbert et Bernays (1939), s'apparentent aux fonctions de choix (2nd ordre) qui elles mêmes se rapprochent des fonctions de Skolem (1er ordre, la quantification sur ces fonctions étant reportée au moment de leur interprétation) — voir par exemple Steedman (2012). Mais les opérateurs de Hilbert ne sont pas comparables avec ces autres formes de quantification : ils incluent les quantificateurs usuels, mais permettent en outre une forme de liage dynamique (comme dans *dynamic predicate logic*) ainsi que des dépendances complexes à la manière des quantificateurs branchants de Henkin. Pour davantage de précision sur les opérateurs de Hilbert, on pourra consulter Slater (2005) ou Avigad et Zach (2008).

Russell (1905) eut le premier l'idée d'introduire un *terme* — un individu — noté $\iota_x P(x)$ comme interprétation logique d'une description définie "le P" où "P" est une propriété, une formule à une variable libre. Que dénote ce terme ? Rien s'il n'existe pas un unique individu tel que $P(x)$, et sinon cet unique individu. Mais chacun sait que le quantificateur "il existe un unique x tel que $P(x)$ " ($\exists! x P(x)$) n'a pas de bonnes propriétés, notamment parce que sa négation, *aucun ou au moins deux* n'a rien de naturel : $\neg \exists! x. P(x) \equiv (\forall x \neg P(x)) \vee (\exists y \exists z (y \neq z) \& P(y) \wedge P(z))$.

Hilbert a donc reformulé ces termes génériques en laissant de côté la condition d'unicité. Il associe un terme $\epsilon_x F(x)$ à toute formule $F(x)$, qui permet d'exprimer la quantification existentielle puisque $F(\epsilon_x F(x)) \equiv \exists x F(x)$. Il introduit aussi son dual $\tau_x F(x)$ qui permet d'exprimer la quantification universelle par $F(\tau_x F(x)) \equiv \forall x F(x)$. Les opérateurs ϵ_x et τ_x lient la variable x dans $F(x)$. Bien sûr, au vu de cette dualité, un seul des deux opérateurs τ et ϵ suffit, si on dispose de la négation. Il est très compliqué d'interpréter ces termes en toute généralité puisqu'on sort de la logique du premier ordre. Les modèles correspondant sont très complexes, voire mal définis (Asser, 1957). En revanche, pour les formules du calcul de Hilbert qui correspondent à des formules habituelles, les modèles usuels fonctionnent et le théorème de complétude est vérifié. En l'absence de modèles simples, définissons la "vérité" de ces formules en termes de déduction, d'autant que les règles de déduction définissant ces opérateurs sont les règles usuelles de la quantification :

- De $F(t)$ où t est n'importe quel terme, on peut déduire $F(\epsilon_x F(x))$ c'est-à-dire $\exists x F(x)$.
- Si l'on a établi $F(x)$ sans rien supposer sur x , on peut en déduire $F(\tau_x F(x))$ c'est-à-dire $\forall x F(x)$

Pour les applications linguistiques, seul ϵ a été utilisé : la quantification existentielle joue un rôle central dans la langue, par exemple la DRT organise le discours autour des quantifications existentielles. L'idée véhiculée par cet ϵ est simplement de construire un terme générique associé au groupe nominal quantifié. Par exemple, pour "un enfant sage" on forme le terme $\epsilon_x.(enfant(x)\&sage(x))$. Selon von Heusinger, pour "l'enfant sage" le terme est quasi identique : $\epsilon_x^1.(enfant(x)\&sage(x))$.⁷ La différence entre ϵ^1 et ϵ n'est qu'une différence d'interprétation : ϵ^1 choisit le plus saillant en contexte tandis que ϵ en choisit un nouveau. Le typage à la Montague de ces opérateurs n'est pas donné. Cependant, ϵ et ϵ^1 sont de type $(e \rightarrow t) \rightarrow e$: ϵ et ϵ^1 produisent un individu (un terme) à partir d'une propriété (ici : $enfant(x)\&sage(x)$), comme le ferait une fonction de choix.

4 Rappels sur le lexique génératif montagovien

Dans (Bassac *et al.*, 2010), nous avons proposé un lexique syntaxique et sémantique qui étend considérablement la sémantique de Montague pour rendre compte de l'adaptation du sens d'un mot au contexte. Ce modèle s'est avéré pertinent pour des questions de sémantique lexicale ou compositionnelle : coprédications possibles ou non, ambiguïté des déverbaux (Real-Coelho et Retoré, 2013), le voyageur fictif (Moot *et al.*, 2011), pluriels (Moot et Retoré, 2011), termes génériques (Retoré, 2012)). Notre modèle est assez proche de (Asher, 2011; Luo, 2011, 2012), mais nous sommes les premiers à aborder la quantification et des déterminants dans un cadre adapté à la sémantique lexicale.

La question initiale qui nous a conduit à utiliser une théorie des types plus sophistiquée que celle de Montague est fort simple : comment rendre compte des restrictions de sélection ? Plus concrètement, comment rejeter les deux premiers exemples et accepter les suivants ?

- (15) * leur **dix** est BON [cf. ex. (19)]
- (16) a. * Une **chaise** ABOIE souvent. [ex. inventé]
- b. Mon **chiot** ABOIE souvent pour m'inciter à jouer avec.
- (17) a. **Barcelone** A BATTU Benfica 2-0. [club]
- b. **Barcelone** (/ *et) A CHOISI DE STRUCTURER LE RÉSEAU ROUTIER de manière à préserver un centre ville piétonnier. [institution]
- (18) a. Mon premier **livre** de cuisine ... Mon livre FÉTICHE à cette époque !
- b. Je l'ai RETROUVÉ, il y a peu, chez ma maman [mon premier livre de cuisine]

En utilisant différents types d'entités et en spécifiant le type d'objet attendu par les prédicats, les compositions sémantiquement impossibles produisent des conflits de types. Le sujet du verbe "aboyer" doit être un chien, ou tout au moins un animal, (16a) un nombre ne saurait être "leur" ni être "bon" (15) etc. L'impossibilité sémantique est matérialisée par l'application d'un prédicat $P^{\xi \rightarrow t}$ présupposant un argument de type ξ (par exemple "animal") à un argument a^α d'un autre type α (par exemple, "meuble") avec $\alpha \neq \xi$: $P^{\xi \rightarrow t} a^\alpha$

7. Les notations de von Heusinger sont source de confusion. Il note η notre ϵ qui correspond au "un" existentiel et qui s'interprète toujours par un nouvel individu, tandis qu'il note ϵ notre ϵ^1 qui correspond à l'article défini "le, la" sans contrainte d'unicité et qui est interprété par l'élément le plus saillant.

On notera qu'il faut parfois relaxer ces contraintes. Dans une discussion sur Internet au sujet d'un prochain match de rugby, l'exemple (15) ci-dessus se trouve :

(19) si leur **dix** est BON ils nous torchent ça c'est sûr

Il faut aussi prévoir que certaines coprédications sont heureuses — (18a et 18b) — et d'autres moins — (17a) et (17b) avec "et" à la place de Barcelone.

Pour traiter tous ces phénomènes, nous avons proposé un lexique sémantique catégoriel où chaque mot se voit associer un λ -terme principal, qui ressemble beaucoup à celui de la sémantique de Montague rappelée ci-dessus, ainsi que des λ -termes optionnels qui permettent de transformer un mot dans l'aspect souhaité, par exemple un numéro en joueur de rugby. En raison du grand nombre de types, il convient de factoriser les opérations sur des termes de types différents et d'avoir des opérations sur des familles de types, et nous nous sommes donc placés dans le λ -calcul du second ordre appelé système F — mais d'autres théories des types comme celle de (Luo, 2012) seraient également possibles. En revanche, notre système se distingue surtout par le caractère lexical et non ontologique des transformations lexicales : celles-ci sont déclenchées par les mots et non par le type des mots. Cela nous semble pleinement justifié par des exemples comme "promotion" et "classe" : les deux désignent des groupes d'élèves, mais seul "classe", peut désigner un lieu (la salle de classe), tandis que "promotion" ne le peut pas. Les coprédications possibles ou impossibles, dans ce système où les mots portent les transformations, sont modélisées en distinguant deux sortes de transformations : les transformations *rigides*, qui imposent de ne référer qu'à cet aspect de l'objet, et les transformations *flexibles* qui permettent de renvoyer à un aspect de l'objet sans exclure les autres aspects dudit objet.

Notre cadre formel, le système F se distingue du lambda calcul simplement typé par l'ajout d'une opération de quantification sur les types dans les termes et les types — cette quantification sur les types joue un rôle déterminant (!) dans notre traitement des déterminants.

Les types sont définis inductivement à partir de types de base :

- Types de base :
 - **t** les valeurs de vérité, **v** les événements,
 - des types constants e_i en grand nombre correspondant aux *différentes sortes d'individus*,
 - des *variables de type*, notées par des lettres grecques (issues d'un ensemble dénombrable P)
 - Lorsque T est un type et α une variable de type, qui peut ou non apparaître dans T , $\Pi\alpha. T$ est un type (dit polymorphe).
 - Lorsque T_1 et T_2 sont des types, $T_1 \rightarrow T_2$ est aussi un type.
- Pour définir les termes, on se donne une infinité dénombrable de variables de chaque type, ainsi que, pour chaque type, des constantes en nombre fini (possiblement aucune) :
- Une variable de type T c'est-à-dire $x : T$ (ce qu'on écrit aussi x^T) est un *terme* de type T .
 - Une constante de type T c'est-à-dire $c : T$ (ce qu'on écrit aussi c^T) est un *terme* de type T .
 - $(f \ \tau)$ est un terme de type U quant τ est de type T et f de type $T \rightarrow U$.
 - $\lambda x^T. \tau$ est un terme de type $T \rightarrow U$ si x est une variable de type T , et τ un terme de type U .
 - $\tau\{U\}$ est un terme de type $T[U/\alpha]$ quand $\tau : \Pi\alpha. T$, et U est un type.
 - $\Lambda\alpha. \tau$ est un terme de type $\Pi\alpha. T$ quand α est une variable de type $\tau : T$ sans occurrence de α dans le type d'une variable libre.

Lorsque les constantes sont celles d'une logique multisorte (connecteurs usuels $\mathbf{t} \rightarrow \mathbf{t} \rightarrow \mathbf{t}$, quantificateurs $\exists, \forall : (e_i \rightarrow \mathbf{t}) \rightarrow \mathbf{t}, \dots$, constantes *Fido* : *ani*, *regarde* : *ani* $\rightarrow \mathbf{e} \rightarrow \mathbf{t}$) ce système est appelé ΛTy_n .

mot	λ -terme principal	λ -termes optionnels	rigide/flexible
<i>Liverpool</i>	$liverpool^T$	$Id_T : T \rightarrow T$ (F) $t_1 : T \rightarrow F$ (R) $t_2 : T \rightarrow P$ (F) $t_3 : T \rightarrow Pl$ (F)	
<i>vaste</i>	$vaste : Pl \rightarrow \mathbf{t}$		
<i>a_voté</i>	$a_voté : P \rightarrow \mathbf{t}$		
<i>a_gagné</i>	$a_gagné : F \rightarrow \mathbf{t}$		

où les types de base sont définis comme suit T (ville), Pl (lieu), P (gens), F (club).

FIGURE 2 – Un exemple de lexique

Les réductions pour λ et Λ sont définies de manière similaire.

- $(\Lambda\alpha.\tau)\{U\}$ se réduit en $\tau[U/\alpha]$ (rappelons que α et U sont des types).
- $((\lambda x^U.\tau^T)^{U \rightarrow T} u^U) : T$ se réduit en $\tau[u/x]$ (réduction habituelle, u est une terme de même type U que la variable x).

La normalisation du système F a une conséquence heureuse pour notre modèle sémantique : si les constantes (du λ -calcul) correspondent au langage L multisorte d'une logique d'ordre n (opérations logiques, prédicats, fonctions et constantes), tout terme normal de type \mathbf{t} correspond à une formule de L .

Donnons l'organisation générale de notre modèle de la sémantique compositionnelle :

le λ -calcul du second ordre, le système F sert à assembler les formules logiques partielles contenues dans le lexique (il remplace le λ -calcul simplement typé utilisé par Montague)

la logique d'ordre supérieur multisorte dans laquelle s'expriment les représentations sémantiques (elle remplace la logique d'ordre supérieur utilisée par Montague, ainsi que ses variantes réifiées du premier ordre : les nombreuses sortes e_i sont les types de base qui gèrent les restrictions de sélection).

Afin d'illustrer l'utilité de la quantification sur les types, donnons un exemple avec une coprédication qui fait intervenir une conjonction polymorphe, donnée en (20c). Cette unique conjonction permet, chaque fois que l'on a deux prédicats $P^{\alpha \rightarrow \mathbf{t}}$, $Q^{\beta \rightarrow \mathbf{t}}$ portant sur des entités de sortes respectives α et β , ainsi que des transformations $f^{\xi \rightarrow \alpha}$ et $g^{\xi \rightarrow \beta}$ du type ξ dans α et dans β de dire que les images d'un objet x^ξ de type ξ ont les propriétés $P^{\alpha \rightarrow \mathbf{t}}$ et $Q^{\beta \rightarrow \mathbf{t}}$.

- (20) a. **Liverpool** est **VASTE** et **A_VOTÉ**.
b. $\&^\Pi\{Pl\}\{P\}(est_vaste)^{Pl \rightarrow \mathbf{t}}(a_voté)^{P \rightarrow \mathbf{t}}\{T\}Liverpool^T(t_3^{T \rightarrow Pl})(t_2^{T \rightarrow P})$
c. $\&^\Pi = \Lambda\alpha\Lambda\beta\lambda P^{\alpha \rightarrow \mathbf{t}}\lambda Q^{\beta \rightarrow \mathbf{t}}\Lambda\xi\lambda x^\xi\lambda f^{\xi \rightarrow \alpha}\lambda g^{\xi \rightarrow \beta}. (and^{\mathbf{t} \rightarrow \mathbf{t} \rightarrow \mathbf{t}}(P(f\ x))(Q(g\ x)))$
d. $(and(est_vaste^{Pl \rightarrow \mathbf{t}}(t_3^{T \rightarrow Pl} Liverpool^T))(a_voté^{P \rightarrow \mathbf{t}}(t_2^{T \rightarrow P} Liverpool^T)))$

Cet exemple s'analyse au moyen des deux transformations ($t_3^{T \rightarrow Pl}$) et ($t_2^{T \rightarrow P}$) celle d'une ville T en un lieu Pl et celle d'une ville en habitants P . Aucune des deux n'étant rigide, on peut les utiliser toutes les deux, et le λ -terme sémantique de la phrase est donné en (20b). On remarquera les spécialisations de types : $\alpha := Pl$, $\beta := P$ et $\xi := T$. Après réduction on obtient comme espéré

(20d). La même situation avec $a_voté$ et $a_gagné$ serait impossible car la transformation d'une ville en club est *rigide*, elle exclut celle de la ville en tant qu'habitants.

5 Des termes typés pour prédicats et déterminants

5.1 Prédicats et types

Un déterminant "classique" s'applique à un prédicat, voire à deux lorsqu'il s'agit d'un quantificateur généralisé, pour donner une proposition. Un opérateur de Hilbert se combine avec un prédicat pour donner un terme. La modélisation du déterminant, avec ou sans opérateurs de Hilbert, est donc intimement liée à celle du prédicat. Dans un système multisorte et typé comme ΛTy_n , il nous faut décider quels sont les types des prédicats : usuellement, un prédicat a pour type $e \rightarrow t$, mais en présence des innombrables types e_i qui se partagent le rôle traditionnellement dévolu à e , que faire ? Faut-il autoriser des prédicats à avoir un domaine autre que e ? Un prédicat comme le nom commun "chat" est-il une propriété du type des "animaux" s'il y en a un, ou est-il une propriété de toutes les entités, propriété qui serait fausse en dehors des "animaux" ?

Cette question est moins embarrassante qu'il n'y paraît car on peut passer d'un choix à un autre. En effet, un prédicat défini sur un type e_i différent de e (le type de toutes les entités), comme $P^{e_i \rightarrow t}$ s'étend en un $\overline{P}^{e \rightarrow t}$ sans difficulté, en disant qu'il est faux en dehors de α . Réciproquement, un prédicat comme *chat* défini sur un type d'entités e_i (par exemple au type *ani* des "animaux") peut être restreint à tout sous type de e_i . Evidemment, un prédicat comme *chat* restreint à un sous ensemble strict de l'ensemble où il est vrai (par exemple au type *siamois*) et ensuite étendu à e puis restreint aux animaux (*ani*) ne redonne pas le prédicat initial, car l'extension est définie uniformément sur tous les types et les prédicats comme étant fausse à l'extérieur du domaine considéré. Ainsi, lorsque β ne contient pas tous les $x : \alpha$ satisfaisant P on a $(P^{\alpha \rightarrow t}|_{\beta})|_{\alpha} \neq P$. Les comportements de la restriction et de l'extension du domaine d'un prédicat se définissent aisément dans un modèle ensembliste.

On peut aussi se demander si un type définit un prédicat. Si *ani* est le type des animaux, y a-t-il un prédicat "être un animal" ? Et si oui, quel est son domaine ? Etant donné un type α il est difficile de dire quel type β contenant α est un bon candidat pour le domaine du prédicat *être de type* α : aussi prendrons nous pour prédicat associé au type α le prédicat $\hat{\alpha}$ de type $e \rightarrow t$

On voit que le type des prédicats, est très lié aux types de base disponibles, qu'aucun chercheur du domaine ne prétend avoir définitivement identifiés. Voici quelques réponses possibles :

- Un seul type e pour toutes les entités, ce qui exclut toute considération lexicale.
- À l'opposé de la solution minimale que nous venons d'évoquer, il y a une solution maximale selon laquelle toute formule à une variable libre définit un type. Il n'est pas sûr qu'un tel système soit bien fondé puisque les formules sont définies au moyen des types.
- Asher (2011) propose d'utiliser un petit nombre de type de base qui correspondraient à des classes ontologiques simples "événement, objet physique, contenu informationnel, humain,..." correspondant aux restrictions de sélection que l'on rencontre dans la langue.
- Luo (2012) propose d'utiliser tous les noms communs.
- Nous n'avons pas d'avis tranché sur la question, mais nous faisons remarquer à la solution précédente, qu'il faut sans doute ajouter aux noms communs des types pour les propositions et

pour les verbes d'action, puisqu'on quantifie aussi sur ce type d'objet :

- (21) a. Elle voudrait qu'il croit en TOUT CE QU'ELLE LUI DIT.
 b. Il a fait TOUT CE QU'IL A PU et il n'a même pas voulu être payé.

5.2 Des termes typés pour les déterminants

Nous proposons que les déterminants indéfinis soit modélisés par une constante ϵ de type $\Pi\alpha. (\alpha \rightarrow \mathbf{t}) \rightarrow \alpha$ — le ϵ de Hilbert adapté au cadre typé et multisorte. Nous voyons donc l'article indéfini comme un ϵ polymorphe, qui se spécialise au type $\{e_i\}$ pour s'appliquer à un prédicat P de type $e_i \rightarrow \mathbf{t}$: il produit un objet du type e_i . Considérons l'exemple suivant, très simple et inventé, afin d'illustrer notre traitement des déterminants (*ani* désigne le types des animaux) :

- (22) a. Un chat dort (sous ta voiture).
 b. terme pour "un" : $\epsilon : \Pi\alpha. ((\alpha \rightarrow \mathbf{t}) \rightarrow \alpha)$
 c. syntaxe : $((un \rightarrow chat) \leftarrow dort)$
 d. sémantique : $dort(un\ chat)$
 e. $(\lambda x. dort^{ani \rightarrow \mathbf{t}}(x))(\epsilon^{\Pi\alpha. ((\alpha \rightarrow \mathbf{t}) \rightarrow \alpha)} chat^{ani \rightarrow \mathbf{t}})$
 f. $(\lambda x. dort(x))(\epsilon^{\Pi\alpha. ((\alpha \rightarrow \mathbf{t}) \rightarrow \alpha)} \{ani\} chat^{ani \rightarrow \mathbf{t}})$
 g. $dort^{ani \rightarrow \mathbf{t}}(\epsilon^{\Pi\alpha. ((\alpha \rightarrow \mathbf{t}) \rightarrow \alpha)} \{ani\} chat^{ani \rightarrow \mathbf{t}}) : \mathbf{t}$
 h. $chat(\epsilon^{\Pi\alpha. ((\alpha \rightarrow \mathbf{t}) \rightarrow \alpha)} \{ani\} chat^{ani \rightarrow \mathbf{t}}) : \mathbf{t}$

La syntaxe fournit un arbre binaire qui indique quel constituant s'applique à l'autre (22c). Comme la sémantique de "chat" est un prédicat qui s'applique aux entités de type "animal", on obtient le λ -terme sémantique (22e). Comme le type du λ -terme sémantique associé à "un" commence par $\Pi\alpha$ (22b), la variable de type α doit s'instancier en *ani*, pour que le terme soit bien typé. On le voit en (22f) avec l'application du terme de "un" au type $ani : \epsilon^{\Pi\alpha. ((\alpha \rightarrow \mathbf{t}) \rightarrow \alpha)} \{ani\}$. Ce terme de type $(ani \rightarrow \mathbf{t}) \rightarrow ani$, est appliqué à *chat* de type $ani \rightarrow \mathbf{t}$, donnant la sémantique de "un chat" qui est de type *ani*. Ce groupe nominal est le sujet du groupe verbal "dort (sous ta voiture)", prédicat qui s'applique à une entité de type "animal". Le terme complet (22f) est bien typé de type \mathbf{t} et il se réduit en (22g) — ce qui, sous des conditions de non vacuité très naturelles, peut se comprendre comme $\exists x : ani \quad dort(x)$.

C'est plutôt satisfaisant, mais rien ne dit que "un chat" ait la propriété d'être un chat ! Le calcul de la sémantique de "un chat" ne produit pas cela. Cependant nous savons $P(\epsilon_x.P(x)) \equiv \exists x P(x)$. Aussi l'énonciation de "un chat dans le sens d'un chat particulier (et non d'un chat générique) nous conduit-elle à ajouter la présupposition $chat(un\ chat)$ (c'est-à-dire $chat(\epsilon_x.chat(x))$) — le λ -terme correspondant est donné en 22h. On notera que *un chat* étant de type "animal" le prédicat "chat" peut effectivement s'y appliquer. Si la présupposition $F(\epsilon_x.F(x))$ est introduite a priori sans avoir rencontré "un F" cela revient à affirmer que la propriété F est satisfaite par au moins un individu. On peut discuter des F pour lesquelles cette présupposition est fondée.

Si "chat" est un type et non une propriété, comme dans Luo (2012), comment faire ? On change le type "chat" en la propriété correspondante \widehat{chat} comme expliqué au paragraphe (5.1), puis on

procède comme ci-dessus.⁸

On peut traiter de la même manière les articles définis, comme le fait von Heusinger : seul le calcul de la référence sera différent. Tandis que l'article indéfini requiert un nouvel élément, l'article défini choisit au contraire un élément déjà présent en contexte. L'approche permet aussi de traiter l'interprétation des pronoms de type E de Evans. Le fait que les termes génériques soient typés n'y change rien. Pour interpréter les anaphores comme le "il" de l'exemple (14) il suffit de recopier le terme sémantique associé à l'antécédent de "il".

La quantification universelle, peu étudiée dans un cadre typé, est extraordinairement simple : elle correspond au terme générique $\tau_x.P(x)$ (c.f. section 3) bien plus facile à interpréter que $\epsilon_x.P(x)$. L'élément $\tau_x.P(x)$ est celui des démonstrations mathématiques : un objet qui, par rapport à F n'a pas de propriété particulière. Ainsi, lorsque $\tau_x.P(x)$ a la propriété P tous les objets l'ont.

6 Implémentation

Le traitement que nous proposons des déterminants et des quantificateurs ne nécessite pas de modifier l'organisation de l'analyseur syntaxique et sémantique du français Grail. L'extension au système F requise par la sémantique lexicale a déjà été testée, du moins sur les parties du lexique dotées d'un typage avec plusieurs sortes, afin de vérifier que ϵ s'instancie convenablement et que la réduction produit les formules quantifiées attendues. La grammaire a été acquise sur corpus, mais à l'heure actuelle nul ne sait comment acquérir automatiquement les lexiques sémantiques convenablement typés que nous utilisons. (Moot, 2010b,a)

Du point de vue syntaxique, les déterminants et quantificateurs ont ici une catégorie plus simple que d'habitude, et surtout ils n'en n'ont qu'une : gn/n suffit alors qu'habituellement il en faut une par position syntaxique. C'est à rapprocher de nos travaux sur l'interprétation sémantique de la grammaire générative (Amblard *et al.*, 2010)

L'implémentation de Moot (2010b) utilise la λ -DRT plutôt que le λ -calcul pour calculer les représentations sémantiques, afin de mieux suivre la structure discursive et aussi de mettre en oeuvre le lien entre opérateurs de Hilbert et liage dynamique des variables existentielles en DRT (von Heusinger, 2004). Techniquement la λ -DRT change peu de choses à notre propos mais aurait nécessité beaucoup de rappels.

7 Conclusion

Ce travail pose à la fois des questions d'analyse sémantique automatique et de logique.

La portée des quantificateurs à la Hilbert doit être discutée ainsi que le lien avec le calcul des prédicats dynamiques. Les opérateurs de Hilbert autorisent des formules sous spécifiées et incluent un liage dynamique : correspondent-ils à ceux couramment utilisés en sémantique ?

8. Une variante : si "chat" est un type et non une propriété, on peut aussi utiliser pour "un" le terme $\epsilon' : \Pi\alpha. \alpha$ (une constante de type \perp ne peut nuire à la cohérence du système). Si on applique cette constante au type "chat" obtient alors "un chat" de type "chat" sans ajouter de présupposition. Comme le fait pertinemment remarquer (Asher, 2011) une déclaration de type $x : T$ est une forme de présupposition, car il est quasi impossible de la nier. On peut appliquer le prédicat "dort" à ce chat, puisque l'inclusion $chat \subset ani$ est une transformation lexicale.

Les pluriels ont aussi un lien avec la quantification, et nous n'en avons pas parlé, en dépit d'un premier travail de Moot et Retoré (2011) dans ce même cadre. Cette question de sémantique est assurément intéressante, elle rejoint des idées anciennes sur les pluriels avec des opérateurs qui gèrent les lectures distributives ou collectives.

Déterminer les types de base est une question importante, surtout en pratique. Peut-être n'y a-t-il pas de réponse en général : leur choix dépend des restrictions de sélection dont on souhaite rendre compte, c'est-à-dire du type d'informations attendues. Par exemple, pour extraire les itinéraires d'un corpus de récits de voyages du XIXe, des types d'entités spatiales et temporelles se dégagent naturellement. (Lefevre *et al.*, 2012)

Notre travail pose également des questions d'acquisition, d'une part des types de base, mais aussi des λ -termes sémantiques, non pour les déterminants qui sont connus ainsi que leur types et termes sémantiques mais pour les autres mots, noms, verbes, adjectifs. Quels sont leurs termes sémantiques ? Avec quels types de base sont-ils écrits ? L'analyseur a besoin de ces informations pour calculer les représentations sémantiques assez fines utilisées ici.

L'interprétation des formules avec ϵ qui ne sont pas équivalentes à des formules usuelles reste mystérieuse. (Slater, 2005; Avigad et Zach, 2008) À ce jour, seule une interprétation très complexe et possiblement erronée a été proposé par Asser (1957) tandis que von Heusinger (2004) a défini une interprétation assez intuitive qui dépend du contexte discursif, mais qui perd l'équivalence avec la quantification usuelle. Peut-on proposer mieux ?

Tant pour la compréhension de la logique sous-jacente que pour l'organisation du modèle d'analyse sémantique automatique, nous souhaiterions mieux comprendre l'interaction entre les types utilisés pour la composition des sens et les prédicats de la logique multi-sorte où s'exprime le sens.

Références

- AMBLARD, M., LECOMTE, A. et RETORÉ, C. (2010). Categorical minimalist grammars : From generative grammar to logical form. *Linguistic Analysis*, 36(1-4):273-306.
- ASHER, N. (2011). *Lexical Meaning in context – a web of words*. Cambridge University press.
- ASSER, G. (1957). Theorie der logischen auswahlfunktionen. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*.
- AVIGAD, J. et ZACH, R. (2008). The epsilon calculus. In ZALTA, E. N., éditeur : *The Stanford Encyclopedia of Philosophy*. Center for the Study of Language and Information. <http://plato.stanford.edu/>.
- BASSAC, C., MERY, B. et RETORÉ, C. (2010). Towards a Type-Theoretical Account of Lexical Semantics. *Journal of Logic Language and Information*, 19(2):229-245. <http://hal.inria.fr/inria-00408308/>.
- BOS, J. (2008). Wide-coverage semantic analysis with boxer. In BOS, J. et DELMONTE, R., éditeurs : *Semantics in Text Processing. STEP 2008 Conference Proceedings*, Research in Computational Semantics, pages 277-286. College Publications.
- CORBLIN, F., COMOROVSKI, I., LACA, B. et BEYSSADE, C. (2004). Generalized quantifiers, dynamic semantics, and french determiners. In CORBLIN, F. et de SWART, H., éditeurs : *Handbook of French Semantic*, chapitre 1, pages 3-22. CSLI Publications.

- EGLI, U. et von HEUSINGER, K. (1995). The epsilon operator and E-type pronouns. In EGLI, U., PAUSE, P. E., SCHWARZE, C., von STECHOW, A. et WIENOLD, G., éditeurs : *Lexical Knowledge in the Organization of Language*, pages 121–141. Benjamins.
- EVANS, G. (1977). Pronouns, quantifiers, and relative clauses (i). *Canadian Journal of Philosophy*, 7(3):467–536.
- GEACH, P. T. (1962). *Reference and generality : an examination of some medieval and modern theories*. Contemporary philosophy. Cornell University Press.
- HILBERT, D. et BERNAYS, P. (1939). *Grundlagen der Mathematik. Bd. 2*. Springer. Traduction française de F. Gaillard, E. Guillaume et M. Guillaume, L'Harmattan, 2001.
- LEFEUVRE, A., MOOT, R., RETORÉ, C. et SANDILLON-REZER, N.-F. (2012). Traitement automatique sur corpus de récits de voyages pyrénéens : Une analyse syntaxique, sémantique et temporelle. In *Traitement Automatique du Langage Naturel, TALN'2012*, volume 2, pages 43–56.
- LUO, Z. (2011). Contextual analysis of word meanings in type-theoretical semantics. In POGODALLA, S. et PROST, J.-P., éditeurs : *LACL*, volume 6736 de *LNCS*, pages 159–174. Springer.
- LUO, Z. (2012). Common nouns as types. In BÉCHET, D. et DIKOVSKY, A. J., éditeurs : *LACL*, volume 7351 de *Lecture Notes in Computer Science*, pages 173–185. Springer.
- MOOT, R. (2010a). Semi-automated extraction of a wide-coverage type-logical grammar for French. In *Proceedings of Traitement Automatique des Langues Naturelles (TALN)*, Montreal.
- MOOT, R. (2010b). Wide-coverage French syntax and semantics using Grail. In *Proceedings of Traitement Automatique des Langues Naturelles (TALN)*, Montreal.
- MOOT, R., PRÉVOT, L. et RETORÉ, C. (2011). Un calcul de termes typés pour la pragmatique lexicale — chemins et voyageurs fictifs dans un corpus de récits de voyages. In *Traitement Automatique du Langage Naturel, TALN 2011*, pages 161–166, Montpellier, France.
- MOOT, R. et RETORÉ, C. (2011). Second order lambda calculus for meaning assembly : on the logical syntax of plurals. In MUSKENS, R., éditeur : *Coconat : Conference on Computing Natural Reasoning*. University of Tilburg. <http://hal.inria.fr/hal-00650644>.
- MOOT, R. et RETORÉ, C. (2012). *The logic of categorial grammars : a deductive account of natural language syntax and semantics*, volume 6850 de *LNCS*. Springer. <http://www.springer.com/computer/theoretical+computer+science/book/978-3-642-31554-1>.
- REAL-COELHO, L.-M. et RETORÉ, C. (2013). A generative montagovian lexicon for polysemous deverbal nouns. In *4th World Congress and School on Universal Logic – Workshop on Logic and linguistics.*, Rio de Janeiro.
- RETORÉ, C. (2012). Variable types for meaning assembly : a logical syntax for generic noun phrases introduced by "most". *Recherches Linguistiques de Vincennes*, 41:83–102. <http://hal.archives-ouvertes.fr/hal-00677312>.
- RUSSELL, B. (1905). On denoting. *Mind*, 56(14):479–493.
- SLATER, B. H. (2005). Epsilon calculi. *The Internet Encyclopedia of Philosophy*. <http://www.iep.utm.edu>.
- STEEDMAN, M. (2012). *Taking Scope : The Natural Semantics of Quantifiers*. MIT Press.
- von HEUSINGER, K. (1997). Definite descriptions and choice functions. In AKAMA, S., éditeur : *Logic, Language and Computation*, pages 61–91. Kluwer.
- von HEUSINGER, K. (2004). Choice functions and the anaphoric semantics of definite nps. *Research on Language and Computation*, 2:309–329.