



HAL
open science

Un traitement sémantique de la polysémie lexicale dans le domaine du dialogue homme-machine finalisé

Evelyne Jacquey, Bertrand Gaiffe, Jean-Marie Pierrel

► To cite this version:

Evelyne Jacquey, Bertrand Gaiffe, Jean-Marie Pierrel. Un traitement sémantique de la polysémie lexicale dans le domaine du dialogue homme-machine finalisé. TALN, Ecole Polytechnique Federale de Lausanne (Martin Rajman) et Universite de Geneve (Eric Wehrli), 2000, Ecole Polytechnique Federale de Lausanne, Suisse, pp.435 - 446. inria-00099064

HAL Id: inria-00099064

<https://inria.hal.science/inria-00099064>

Submitted on 26 Sep 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Un traitement sémantique de la polysémie lexicale dans le domaine du dialogue homme-machine finalisé

Jacquey Evelyne, Gaiffe Bertrand et Pierrel Jean-Marie

email : Evelyne.Jacquey@loria.fr, Bertrand.Gaiffe@loria.fr, Jean-Marie.Pierrel@loria.fr

tel : 03 83 59 20 34, fax : 03 83 41 30 79

adresse : Batiment LORIA, Campus Scientifique, BP 239, 54506 Vandœuvre les Nancy

Résumé

L'objet de cet article est de proposer un traitement lexical des noms polysémiques qui admettent des coprédications avec ou sans reprises pronominales. Ces coprédications peuvent combiner des prédicats véhiculant des restrictions sélectionnelles différentes et incompatibles à première vue. Ces données semblent donc remettre en cause l'hypothèse qu'une coprédication, par le biais d'une coordination ou d'une reprise pronominale, ne modifie pas le type de l'élément qui subit la coprédication. Parmi un certain nombre de travaux existants dans ce domaine, (Godard & Jayez, 1996) et (Pustejovsky, 1994), (Pustejovsky, 1995) entre autres, (Asher & Pustejovsky, 2000) proposent un traitement qui vérifie l'hypothèse d'un typage constant des arguments par leur prédicat sans pour autant interdire ces phénomènes de coprédication. Nous montrerons que ces travaux constituent une alternative à l'hypothèse de (Godard & Jayez, 1996). De plus, dans la perspective de l'élaboration d'une interface de dialogue de commandes pour guider les tâches accomplies par une application préexistante, nous montrerons que le traitement proposé par (Asher & Pustejovsky, 2000) fournit des représentations sémantiques adaptables à deux types possibles de modèles conceptuels.

Mots-clés : polysémie nominale, restrictions sélectionnelles des prédicats, inférence de types, typage complexe, traitement sémantique, modèle d'application.

1. Introduction et problématique

L'objet de cet article est de proposer un traitement d'une classe de noms particuliers qui est caractérisée par deux propriétés.

- Les noms de cette classe sont *polysémiques*. Un mot polysémique donne accès hors-contexte à plusieurs sens possibles pour lesquels on peut déterminer une relation stable. Si chacun de ces sens est associé à un type distinct, on peut donc définir un mot polysémique comme une liste de sens de ce mot, chacun étant caractérisé par un type.
- Ces noms admettent des *coprédications avec ou sans reprises pronominales* et ces coprédications ont la caractéristique de combiner des prédicats véhiculant des restrictions sélectionnelles différentes et non compatibles.

Un exemple prototypique, en français, est le nom *livre*. Ce nom est au moins ambigu entre un sens matériel et un sens informationnel, le premier sens dénotant le support de l'objet correspondant au second sens. En termes de types, les groupes nominaux (NPs), qui ont pour tête le nom *livre*, peuvent être de type *p* (objet-matériel) (1), de type *i* (objet-informationnel) (2b) ou des deux types à la fois (2a)¹.

- (1) a. Son livre **est sur l'étagère**.
b. Ce livre **est lourd**.
- (2) a. Pierre **a lu** ce livre hier.
b. Ce livre **est très intéressant**.

Les coprédications où figurent ces NPs peuvent combiner des restrictions sélectionnelles qui semblent, à première vue, incompatibles.

- (3) a. Ce livre **est lourd, mais est très intéressant**.
b. Pierre **a volé et lu** ce livre.
c. Pierre **a trouvé** ce livre **sur l'étagère**. Il l'**a lu** hier.

Ces données semblent entrer en contradiction avec trois hypothèses fondatrices de nombreux travaux sur le lexique².

- **(H1)** Les restrictions sélectionnelles des prédicats permettent de typer leurs arguments.
- **(H2)** Ce typage par les prédicats est un typage fort : typage constant sauf coercion particulière et nombre restreint de types affectés³.
- **(H3)** Le typage d'un groupe nominal peut être déterminé en examinant les contextes sélectifs dans lesquels ce NP peut être repris par un pronom. Si cette coprédication avec reprise pronominale est possible, on conclut que le NP est du même type que le pronom, type qui est déterminé par le biais des restrictions sélectionnelles du prédicat principal dans le contexte sélectif utilisé.

1. Pour une description détaillée des critères qui permettent de typer les NPs, voir les travaux de (Godard & Jayez, 1996).

2. On peut citer parmi d'autres les travaux de (Grimshaw, 1990), (Levin, 1993), (Tenny, 1994), etc.

3. Cette seconde hypothèse permet notamment d'éviter l'écueil d'une très grande complexité et d'une forte redondance du lexique.

Ces hypothèses s'appuient sur deux types d'arguments : des arguments liés à l'usage de la langue et d'autres liés à la mise en œuvre du traitement automatique des langues.

La plupart des locuteurs ont un certain nombre d'intuitions claires sur le type de certains noms, comme *chaise* par exemple qui est habituellement considéré comme désignant un objet matériel, ou sur les restrictions sélectionnelles de certains prédicats, comme *être sur l'étagère* qui est souvent considéré comme sélectionnant un objet matériel. Lorsque l'on veut mettre en œuvre la construction d'un lexique, il semble intéressant de s'appuyer au moins en partie sur ces intuitions courantes. L'hypothèse **H2** permet alors de garder l'acquis de ces intuitions courantes tandis que l'hypothèse **H3** est utilisée pour étendre le lexique. Enfin, dans le cadre des interfaces de dialogue de commandes destinées à guider les actions d'une application, on peut supposer que les objets et les tâches de cette application sont fortement typées. Il serait donc intéressant d'obtenir des représentations sémantiques qui conservent un typage constant des arguments par leur prédicat.

Dans cet article, le but est de montrer qu'il est possible de rendre compte lexicalement des propriétés qui caractérisent des noms comme *livre*, c'est-à-dire de résoudre les conflits de types dûs à la combinaison de restrictions sélectionnelles, incompatibles à première vue, sans pour autant remettre en cause ces hypothèses.

2. Le traitement proposé dans (Godard & Jayez, 1996)

(Godard & Jayez, 1996) font deux hypothèses. D'une part, les noms polysémiques, considérés ici comme des noms multitypés, sont associés à un ensemble de types qu'ils appellent "la combinatoire de types". D'autre part, un NP hérite des types du nom qui en est la tête lexicale.

Dans le cas de *livre*, les deux types correspondant aux deux sens accessibles sont considérés par (Godard & Jayez, 1996) comme "compatibles", c'est-à-dire combinables. Ainsi, *livre* est caractérisé par la combinatoire de types, $\{p, i, p \& i\}$. Les possibilités de coprédication exhibées par les NPs dont *livre* est la tête peuvent alors être expliquées par le fait que chacun des prédicats combinés sélectionne un type particulier dans cette combinatoire de types.

L'exemple de coprédication *Un homme a volé et lu ce livre*, sachant que *voler* demande un argument de type p et *lire* un argument de type $p \& i$, se représente sémantiquement par la forme logique suivante⁴.

$$(4) \exists x \exists y (Homme(y) \wedge Livre(x) \wedge Voler(y, x) \wedge Lire(y, x))$$

avec $x : \{p, i, p \& i\}$ et $y : \text{humain}$

Ce traitement permet d'expliquer les phénomènes de coprédication lorsque ceux-ci ne tiennent pas compte de l'incompatibilité entre les restrictions sélectionnelles des prédicats combinés. Cependant, l'hypothèse **H2**, c'est-à-dire l'hypothèse du typage fort sur les arguments d'un prédicat, n'est respectée que relativement à la combinatoire de types. Selon **H2**, *voler* sélectionne un argument de type p uniquement. Or, comme *voler* porte sur $x : \{p, i, p \& i\}$ grâce à la combinatoire de types, tout se passe comme si, dans la forme logique, le verbe *voler* était susceptible de sélectionner les trois types fournis.

4. Dans l'ensemble des formes logiques figurant dans cet article, nous avons sous-spécifié la représentation du temps et des éventualités pour éviter de compliquer inutilement les représentations.

3. La notion de “Dot-object”, (Pustejovsky, 1995)

Observant que les différents sens véhiculés par un mot polysémique⁵ sont également accessibles hors-contexte et que ceux-ci peuvent être explicités par une relation synchronique, (Pustejovsky, 1995) introduit la notion de “Dot-object” qui lui permet de proposer un traitement basé sur une notion étendue de la compositionnalité.

La notion de “Dot-object” peut se voir comme un type complexe, noté $\alpha \bullet \beta$, mais d’un genre différent des types complexes obtenus par unification (cf. (Carpenter, 1992) ou (Copestake *et al.*, 1993)). Dans ce cas, les deux types simples ne sont pas unifiables par le biais de leur plus petit surtype commun. Cette notion sert à modéliser deux aspects des noms polysémiques. Un nom de type pointé $\alpha \bullet \beta$ est en fait ambigu entre trois sens, les deux premiers correspondent aux types simples α et β et le troisième au type pointé. Ces trois sens forment un paradigme lexical conceptuel (“Lexical Conceptual Paradigm” ou LCP⁶). Le LCP de *livre*, qui est associé au type pointé $p \bullet i$, s’écrit alors `phys_obj•info_lcp = {p•i, p, i}`. Ce paradigme lexical conceptuel correspond exactement à la combinatoire des types prônée par (Godard & Jayez, 1996). Cependant, l’hypothèse de Pustejovsky introduit deux différences sensibles par rapport à l’explication de (Godard & Jayez, 1996) :

- Dans les entrées lexicales de noms de type pointé $\alpha \bullet \beta$, chacun des types simples combinés est représenté par deux variables distinctes, $x : \alpha$ et $y : \beta$, et le lien que ces types entretiennent est explicité par la valeur du rôle FORMAL⁷ que l’on peut schématiser par une relation telle que $R(x,y)$. Dans le cas de *livre*, R est une relation de type `contenant / contenu`. Ainsi, selon (Pustejovsky, 1995), les deux variables de types simples, `p` et `i`, sont liées dans le rôle FORMAL par le prédicat statique $Hold(e^s, x, y)$ où la variable e^s désigne une éventualité de type `état`.
- Dans la forme logique associée à l’exemple *Un homme a volé et lu ce livre, voler* porte cette fois sur un argument de type `p` et *lire* sur un argument de type `p•i`. On aurait ainsi⁸ :

$$(5) \exists x \exists y \exists z (Homme(x) \wedge Livre(y.z) \wedge Voler(x,y) \wedge Lire(x,y.z) \wedge Hold(y,z))$$

avec $x : \text{humain}, y : p, z : i$ et $y.z : p \bullet i$

Les formes logiques obtenues fournissent une représentation sémantique satisfaisante de la coprédication, et cela tout en vérifiant les trois hypothèses **H1**, **H2**, **H3**. Cependant, c’est un des reproches courants qui sont faits au lexique génératif, de nombreuses difficultés apparaissent dans une mise en œuvre effective des mécanismes proposés⁹. Récemment, une mise en œuvre

5. Sa définition de la polysémie est très proche de celle qui est donnée ici. Elle est basée sur les deux types d’ambiguïtés lexicales décrites par (Weinreich, 1964) : les ambiguïtés contrastives, qui apparaissent entre des mots entretenant une relation d’homonymie, et les ambiguïtés complémentaires, qui apparaissent entre les différents sens d’un même mot, ces sens entretenant alors une relation de polysémie.

6. Plus précisément, la notion de LCP, qui a été introduite par (Pustejovsky & Anick, 1988), correspond à l’ensemble structuré des sens possibles d’un même mot polysémique.

7. Ce rôle de la structure de qualia décrit les critères qui caractérisent la classe sémantique du terme dont on construit l’entrée lexicale.

8. Les éventualités n’étant pas représentées, il y a un décalage entre la valeur du rôle FORMAL et la manière dont le prédicat *Hold* est exprimé dans la forme logique (5).

9. L’approche du lexique génératif a donné lieu à de nombreux articles critiques. La revue de (Habert & Nazarenko, 1996) fournit une vision synthétique des avantages et des inconvénients de cette approche. On citera aussi les critiques faites par (Gayral, 1998) concernant une généralisation possible des mécanismes de coercion de types ou encore de l’utilisation des informations fournies dans les structures de qualia.

formelle de cette approche a été proposée dans (Asher & Pustejovsky, 2000)¹⁰.

4. Modélisation de la notion de “Dot-Object”, (Asher & Pustejovsky, 2000)

La finalité de la notion de “Dot-object” reste la même : rendre compte des phénomènes de coprédication ne respectant pas la compatibilité des restrictions sélectionnelles véhiculées par les prédicats mis en jeu, mais sans pour autant remettre en cause les trois hypothèses **H1**, **H2**, **H3**.

4.1. Trois éléments nouveaux pour définir la notion de “Dot-object”

Défendant l’idée qu’un type pointé, donc le type d’un “dot-object”, est le résultat d’une combinaison de types particulière, (Asher & Pustejovsky, 2000) introduisent un nouveau constructeur de type, noté \bullet , et qui se traduit par le principe suivant : “Si α et β sont des types, alors $\alpha \bullet \beta$ est un type”.

A la différence de (Pustejovsky, 1995), (Asher & Pustejovsky, 2000) ne définissent pas de relation de sous-typage entre le type pointé et les surtypes de chacun des types simples combinés¹¹.

Convaincus que la représentation conceptuelle que l’on a du monde (encore appelée la “métaphysique de sens commun”) et le lexique doivent être dissociés sans nier leur interdépendance, la relation introduite dans le rôle FORMAL du lexique génératif est maintenant décrite à l’aide de deux relations qui n’entrent pas dans la définition des entrées lexicales mais “traduisent” conceptuellement le type pointé.

- (6) a. $book_Elaboration_Phys(x,y) \wedge x : p \bullet i \rightarrow y : p$
 b. $book_Elaboration_Info(x,y) \wedge x : p \bullet i \rightarrow y : i$

Ces deux règles signifient qu’à partir de tout objet complexe x de type $p \bullet i$, la conjonction avec l’une ou l’autre des relations $book_Elaboration$ permet d’extraire l’un des aspects de cet objet complexe : son aspect physique y de type p avec la relation $book_Elaboration_Phys(x,y)$ en (6a) ou bien son aspect informationnel en (6b) avec la relation $book_Elaboration_Info(x,y)$ qui associe le type i à la variable y .

Enfin, lorsque les restrictions sélectionnelles du prédicat et le type de l’argument rencontré sont différents, on se trouve face à un conflit de types qu’il faut résoudre pour pouvoir β -réduire. La résolution des conflits de types induits par des types pointés nécessite des règles particulières d’inférence de type, car on ne peut s’appuyer sur une hiérarchie de types.

Les règles définies dans (Asher & Pustejovsky, 2000) partagent un mécanisme général :

- Les conflits sont indiqués au niveau des variables elles-mêmes. Les assignations de types, pour chacune des variables, sont indiquées dans un contexte (la liste entre $\langle \dots \rangle$).

$$(7) \exists t \text{ Livre}(t) \wedge \lambda v \lambda x (\text{Voler}(v,x))[t], \langle t : p \bullet i, v : h, x : p \rangle$$

10. (Jabez & Godard, 1995) ont proposé une mise en œuvre critique de certaines hypothèses du lexique génératif dans le cadre de la théorie HPSG, mais celle-ci est davantage centrée sur les phénomènes de métonymie.

11. Voir la relation “Dot object subtyping”, dans (Pustejovsky, 1995):150.

- Les règles d’inférence de type, associées aux types pointés, utilisent toutes une règle conceptuelle. Le choix de la règle appropriée dépend de deux critères :
 1. l’existence d’une entrée lexicale du type pointé impliqué dans le conflit,
 2. le type simple impliqué dans le conflit, c’est-à-dire α ou β .
 Concernant les NPs dont *livre* est la tête, on utilisera (6a) lorsque le type simple est p , et (6b), avec le type i .
- Chaque application de règle consiste à introduire une variable existentielle qui correspond à l’aspect manquant dans la formule initiale. Cette variable est ensuite liée par l’intégration de la règle conceptuelle appropriée.
De plus, les règles utilisées se répartissent en deux ensembles selon le type d’opération appliquée au type pointé.

4.2. Deux types de règles selon l’opération subie par le type pointé

– Le type pointé est exploité.

On donne ici l’exemple d’une règle “Dot-exploitation-left” ou “DEL”, qui résout un conflit entre la variable λ -abstraite x de type α et un terme $[t]$ de type $\alpha \bullet \beta$.

$$\frac{\lambda x \phi[t], c(x : \alpha, t : \alpha \bullet \beta) \quad (\lambda P S(v,w)(\phi) \wedge v : \alpha \bullet \beta) \rightarrow w : \alpha}{(\lambda y \lambda \vec{u} \exists x (\lambda P S(y,x)(\phi)[\vec{u}] \wedge \phi[\vec{u}]))[t], c + y : \alpha \bullet \beta}$$

Le conflit de type, tel qu’il apparaît en (7), est résolu en utilisant une règle DEL et la règle conceptuelle (6a). On sait qu’on doit résoudre le conflit entre x et t ¹². Le terme $[t]$ est de type $p \bullet i$, donc on introduit un y de même type, on réintroduit existentiellement x et on intègre la règle conceptuelle. On obtient ainsi une nouvelle forme logique.

$$(8) \exists t \text{ Livre}(t) \wedge \lambda y (\lambda v \exists x \text{ Voler}(v,x) \wedge \text{book_Elaboration_Phys}(y,x))[t], \\ < t, y : p \bullet i, v : h, x : p >$$

Ce qui après β -réduction donne,

$$(9) \exists t \exists x \text{ Livre}(t) \wedge \lambda v \text{ Voler}(v,x) \wedge \text{book_Elaboration_Phys}(t,x) \\ < t : p \bullet i, v : h, x : p >$$

– Le type pointé est introduit.

On donne l’exemple de la règle “Dot-introduction-left” ou “DIL”, qui résout un conflit entre un x de type $\alpha \bullet \beta$ et un terme $[t]$ de type α .

$$\frac{\Psi[\lambda x \phi[t]], c(x : \alpha \bullet \beta, t : \alpha) \quad (S(y,z)(\phi) \wedge y : \alpha \bullet \beta) \rightarrow z : \alpha}{\Psi_x^{\exists} (S(t,y) \wedge \lambda x \phi[t]), c * (t : \alpha \bullet \beta)}$$

Si on reprend l’exemple de coprédication utilisé jusqu’à maintenant, *Un homme a volé et lu ce livre*, on rencontre un premier conflit dans la construction de la forme logique associée au groupe de prédicats *a volé et lu*¹³.

$$(10) \lambda x \lambda v \text{ Voler}(v,x) \wedge \lambda u (\text{Lire}(v,u))[x], < x : p, v : h, u : p \bullet i >$$

12. L’analyse syntaxique fournit l’information que le NP dont *livre* est la tête correspond à la deuxième position argumentale de *Voler*.

13. En (10), on ne montre que l’étape où il faut résoudre un conflit de type, donc ce qui concerne la β -réduction sur la seconde position argumentale de *Lire*. La β -réduction sur sa première position argumentale ne pose pas de problème particulier car les prédicats *Voler* et *Lire* demandent tous deux un premier argument de type h (humain).

Dans ce cas, on utilise une règle DIL et la règle conceptuelle (6a). Avec $[\psi = [\lambda x \lambda v (Voler(v, x))]]$ et $[\phi[x] = [\lambda u Lire(v, u)[x]]]$, on obtient une nouvelle forme logique dans laquelle y a pris la place de x et son type initial, sauf dans ϕ , et x est retypé en $p \bullet i$.

$$(11) \exists y \lambda x \lambda v Voler(v, y) \wedge book_Elaboration_Phys(x, y) \wedge \lambda u (Lire(v, u))[x] \\ < x : p \bullet i, v : h, u : p \bullet i, y : p >$$

Ce qui, après β -réduction donne,

$$(12) \exists y \lambda x \lambda v Voler(v, y) \wedge Lire(v, x) \wedge book_Elaboration_Phys(x, y) \\ < x : p \bullet i, v : h, y : p >$$

Comme on peut l'observer, la forme logique obtenue en (12) fournit une représentation sémantique satisfaisante pour le groupe *a volé et lu* et dans le même temps, vérifie les trois hypothèses **H1**, **H2**, **H3**.

Pour construire ensuite l'ensemble du VP *a volé et lu ce livre*, il suffit d'assembler la forme logique de *ce livre* et la forme logique de *a volé et lu* en (12). Cette dernière étape ne pose pas de problème car les types demandés par *livre* et la variable λ -abstraite x en (12) sont les mêmes. On obtient ainsi :

$$(13) \exists x \exists y \lambda v Livre(x) \wedge Voler(v, y) \wedge Lire(v, x) \wedge book_Elaboration_Phys(x, y) \\ < x : p \bullet i, y : p, v : h >$$

5. Ce traitement dans la perspective de l'élaboration d'une interface de dialogue de commandes

5.1. Bilan des travaux existants

La notion de "Dot-object" fournit une explication satisfaisante des phénomènes de coprédication et, de plus, vérifie les hypothèses **H1**, **H2**, **H3**.

On peut faire brièvement deux observations sur la formalisation de la notion de "Dot-object" telle qu'elle est présentée dans (Asher & Pustejovsky, 2000).

- Les représentations sémantiques obtenues peuvent s'adapter à deux types possibles de modèles conceptuels, l'un contenant des objets multitypés et l'autre ne contenant que des objets de types simples.
- Les règles d'inférence de type donnent cependant accès à une procédure trop puissante.

5.2. Adaptabilité de ce type de représentation sémantique à un modèle conceptuel

Comme nous l'avons dit plus haut, les représentations sémantiques sous forme logique, que nous obtenons en appliquant les règles d'inférence de type, peuvent s'adapter à deux types de modèles conceptuels, le premier contenant des objets complexes associés aux variables de type pointé, le second ne contenant que des objets de type simple, y compris pour les variables de type pointé. Pour illustrer notre propos, imaginons une application permettant de manipuler des objets graphiques. Pour interpréter un ordre tel que *Fais la reproduction de ce triangle et mets-la ensuite en haut de l'écran* en utilisant la méthode proposée dans (Asher & Pustejovsky, 2000), *reproduction* est associé à un type complexe $process \bullet figure$ et ce type complexe

s'exprime par deux règles conceptuelles, les règles (14a) et (14b), qui sont construites sur le modèle des règles associées au type pointé de *livre* (6a-6b).

- (14) a. $Reproduction_Elaboration_Proc(x,y) \wedge x : proc \bullet fig \rightarrow y : proc$
 b. $Reproduction_Elaboration_Fig(x,y) \wedge x : proc \bullet fig \rightarrow y : fig$

Le prédicat *faire* demande ici un objet de type `process` et *mettre+Loc*, un objet de type `figure`. Le contenu sémantique de l'ordre à interpréter peut être représenté par la forme logique suivante.

- (15) $\exists v \exists y \exists z (Reproduction(v) \wedge Faire(y) \wedge Reproduction_Elaboration_Proc(v,y) \wedge Mettre-dans - Loc(z) \wedge Reproduction_Elaboration_Fig(v,z))$
 avec $v : proc \bullet fig$, $y : proc$, et $z : fig$

Pour accomplir effectivement cet ordre, il faut ensuite “enrichir” cette forme logique, enrichissement qui dépend du modèle conceptuel choisi pour l'application.

- **Le modèle de l'application contient des objets complexes de type `process`•`figure`.**

Dans ce cas, les trois variables v, y et z , respectivement associées aux types `process`•`figure`, `process` et `figure` doivent pointer sur un même objet complexe de l'application. Enrichir la forme logique (15) consiste alors à projeter les deux relations conceptuelles impliquées dans (14a) et (14b) sur la relation d'identité, ce qui conduit à poser $v = y = z$.

- **Le modèle de l'application contient uniquement des objets simples, de type `process` ou `figure`, types qui sont reliés par des pointeurs `PT_PROC` et `PT_FIG`.**

Dans un tel modèle, enrichir la forme logique conduit à projeter les deux variables de type simple y et z sur des objets simples de l'application de types respectivement compatibles. Pour cela, les relations conceptuelles figurant dans (14a) et (14b) sont projetées respectivement sur les deux pointeurs `PT_PROC` et `PT_FIG`. Il faut noter ici que la variable de type complexe v n'a pas de référent.

5.3. Surpuissance des règles d'inférence de type

Dans une formule courante $\lambda x \phi[t]$, un conflit entre la variable λ -abstraite x , de type pointé $\alpha \bullet \beta$, et le terme $[t]$, de type simple α ou β , laisse le choix entre deux résolutions possibles, utiliser une règle de “Dot-exploitation” ou une règle de “Dot-introduction”. Selon les cas, les deux modes de résolution possibles conduisent à des formes logiques identiques, et donc redondantes, ou bien à des formes logiques différentes.

Dans le cas de l'exemple simple *Un homme a volé ce livre*, on peut résoudre le conflit comme nous l'avons montré en (8-9), donc avec une règle de “Dot-exploitation”. Pour résoudre ce conflit, une règle de “Dot-introduction” pourrait aussi être utilisée. Ces deux types de résolution fournissent une forme logique identique à la forme logique du VP *a volé ce livre* en (9).

Pour une phrase telle que *Jean a volé tous les livres de la bibliothèque* en revanche, les différents types de résolutions possibles conduisent à des résultats différents sur le plan de la

quantification.

Le quantificateur *tous* est associé au λ -terme $\forall x \lambda P \lambda Q (P(x) \rightarrow Q(x))$. Selon le type de résolution choisi, le quantificateur universel \forall porte sur une variable de type p ou $p \bullet i$. Or, la forme logique attendue est celle où \forall porte sur une variable de type p , c'est-à-dire l'interprétation dans laquelle l'interlocuteur comprend que *Jean* a volé tous les objets de la bibliothèque qui ressemblaient physiquement à un livre.

Par conséquent, il serait intéressant de définir des contraintes sur l'application des règles d'inférence de type pour réduire le nombre de formes logiques potentiellement produites. Le but serait d'une part de réduire le nombre de formes logiques identiques et d'autre part, d'éviter la construction de formes logiques correspondant à des interprétations douteuses. Deux pistes semblent envisageables.

- *Contrôle de la tête du constituant*

Cette restriction semble être appliquée dans (Asher & Pustejovsky, 2000) sans qu'elle soit explicitement formulée. Elle nous semble néanmoins fournir un moyen de restreindre l'application des règles et nous en proposons donc une formulation possible¹⁴.

“L'application de toute règle d'inférence doit tenir compte du type de la tête sémantique du constituant dont on est en train de construire la forme logique”

Dans la mesure où le traitement sémantique proposé repose sur l'hypothèse du typage des arguments par le biais des restrictions sélectionnelles de leur prédicat, il semble légitime de guider l'application des règles selon ces restrictions sélectionnelles. Plus généralement, si l'on suppose qu'une construction attributive se réécrit minimalement ((Det N) être Adj), au cours de l'assemblage des termes correspondant à chacun des éléments mis en jeu, on suivrait l'ordre suivant : [Det appliqué à N \rightarrow NP], puis, [NP appliqué à être Adj \rightarrow S]. La tête du constituant en cours de construction est donc le terme qui subit l'application. On obtiendrait ainsi les formes logiques attendues pour des phrases comme *Tous les livres sont lourds*, avec une quantification universelle sur une variable de type p , et *Tous les livres sont intéressants*, avec une quantification universelle sur une variable de type i . Cependant, pour un exemple de copredication tel que *Tous ces livres sont abîmés mais ils sont intéressants*, la variable de discours accessible pour le pronom *ils* est de type p et la construction *sont intéressants* demande une variable de type i . Un conflit de cette nature conduit (Asher & Pustejovsky, 2000) à proposer l'utilisation d'une règle de “Dot-introduction” sur la formule courante et sur le terme associé à *sont intéressants*. Or une telle procédure conduit à une forme logique dans laquelle la variable introduite de type $p \bullet i$ est quantifiée universellement, ce qui va à l'encontre du respect du type de la tête.

14. Nous avons fait le choix de la tête sémantique plutôt que la tête syntaxique en nous appuyant notamment sur les travaux de (Pollard & Sag, 1987) et (Pollard & Sag, 1994). De manière très abrégée, un aspect de ces travaux concerne la manière dont le contenu sémantique est propagé au long de l'analyse d'un constituant et de la reconstruction parallèle de son contenu sémantique. Le contenu sémantique propagé est celui de la tête sémantique du constituant alors que la tête syntaxique de celui-ci est toujours sa tête lexicale. (Pollard & Sag, 1987) et (Pollard & Sag, 1994) ont montré que, lorsque la tête lexicale du constituant n'est pas spécifiée par un modifieur (adjectif, adverbe, etc.), c'est le contenu sémantique de cette tête lexicale qui est propagé. Dans ce cas, la tête syntaxique et la tête sémantique du constituant correspondent au même élément. En revanche, lorsque la tête lexicale est spécifiée par un modifieur, c'est alors le contenu sémantique de ce modifieur qui est propagé sachant que le contenu sémantique de la tête lexicale est intégré dans celui du modifieur. Dans ce second cas, la tête sémantique du constituant est le modifieur.

- *Contrôle selon le nombre de conflits résolus*

L'autre piste envisageable est de privilégier le mode de résolution qui résout le nombre maximum de conflits de type dans l'ensemble de la forme logique non β -réduite. Dans le cas de la phrase *Jean a volé tous les livres de la bibliothèque*, appliquer cette contrainte conduit à privilégier l'application d'une règle de "Dot-exploitation" à partir du type pointé de *livre*. On obtient donc une formule intermédiaire quantifiée universellement sur une variable de type p , formule qui peut s'appliquer ainsi sans difficulté au terme représentant *a volé*. En conclusion, cette seconde contrainte semble efficace et respecte le typage de la tête. Dans le cas de *Jean a lu tous les livres de la bibliothèque*, appliquer cette contrainte consiste à appliquer une règle de "Dot-introduction" sur *bibliothèque*. L'efficacité de la méthode reste vraie, mais il n'en est pas de même pour le respect du typage de la tête. Dans le constituant *livres de la bibliothèque*, la tête est la préposition *de* qui, dans ce cas, requiert un argument de type p pour tout objet se trouvant dans la bibliothèque. En appliquant une règle de "Dot-introduction", on ne respecte pas cette contrainte de type.

6. Conclusion et perspectives

Nous avons montré qu'en s'appuyant sur le traitement proposé par (Asher & Pustejovsky, 2000), il est possible de traiter au niveau sémantique les conflits de types dûs à la présence de noms polysémiques dans un certain nombre de coprédications qui se révélaient *a priori* problématiques.

Cependant, comme l'observent (Godard & Jayez, 1996), tous les noms polysémiques n'admettent pas ce type de coprédication avec le même degré de liberté.

- Certains noms ne l'admettent pas du tout ou très rarement. C'est le cas de déverbaux de création tels que *construction*. L'explication fournie repose sur le fait les types *evt* et *obj-mat* seraient exclusifs. Cependant, même s'ils ne permettent pas la coprédication, ils satisfont au moins une des propriétés des mots polysémiques : il existe un lien entre leurs deux sens possibles, l'objet matériel est le résultat de l'événement. Comme le montre (16), *Marie* est préférentiellement interprétée comme référant à l'agent du processus de construction, plutôt qu'à la détentrice de l'objet matériel résultant.

(16) La jolie construction de Marie est sur l'étagère.

- D'autres noms, qui pourtant sont polysémiques, admettent la coprédication dans certains cas seulement, c'est le cas de déverbaux tels que *évaluation*, *description*, etc. Ce deuxième cas de figure s'explique par le passage par un type intermédiaire, qui serait une extension très abstraite du type *obj-info*, le type *organisation*. Concernant ce dernier cas notamment, il nous semble intéressant d'aménager le traitement sémantique proposé par (Asher & Pustejovsky, 2000) de manière à rendre compte de ces phénomènes.

Un autre point nous semble à étudier. D'une part, les règles d'inférence pour les types pointés nécessitent d'être contraintes davantage, et notamment concernant la combinaison de plusieurs types de contraintes. Par ailleurs, l'accessibilité des variables introduites existentiellement par le biais des règles d'inférence sont un autre axe de recherche. La question est ici de savoir si ces variables sont des référents de discours accessibles ou non. Face à une anaphore telle que *Ne lui offre pas ce livre, il l'a déjà*, on aimerait que non, car les deux objets "livre" se distinguent

nécessairement du point de vue physique. *A contrario*, dans le cas de *Ne lui offre pas ce livre, il est déchiré*, on aimerait que oui, car cette fois, on parle bien du même objet physique.

Références

- ASHER N. & PUSTEJOVSKY J. (2000). The metaphysics of words in context. submitted to Journal of Logic, Language and Information. <http://www.dla.utexas.edu/depts/philosophy/faculty/asher/papers/MWC.ps>.
- CARPENTER B. (1992). Typed feature structures. *Computational Linguistics*, **18**(2).
- COPESTAKE A., SANFILIPPO A., BRISCOE T. & DE PAIVA V. (1993). The ACQUILEX LKB: An introduction. In T. BRISCOE, V. DE PAIVA & A. COPESTAKE, Eds., *Inheritance, Defaults and the Lexicon*. Cambridge University Press.
- GAYRAL F. (1998). Créativité du sens en contexte et hypothèse de compositionnalité. *T.A.L.*, **39**(1), 67–98.
- GODARD D. & JAYEZ J. (1996). Types nominaux et anaphores ; le cas des objets et des vnements. In *Anaphores temporelles et (in-)cohérence*. Cahiers Chronos.
- GRIMSHAW J. (1990). *Argument Structure*. MIT Press.
- HABERT B. & NAZARENKO A. (1996). Revue de "the generative lexicon" de Pustejovsky(95). *T.A.L.*, **37**(2), 159–163.
- JAYEZ J. & GODARD D. (1995). Principles as lexical methods. In *Proceedings of the AAAI workshop on representation and acquisition of lexical knowledge*.
- LEVIN B. (1993). *Towards a Lexical Organization of English Verbs*. University of Chicago Press.
- POLLARD C. & SAG I. (1987). *Information-Based Syntax and Semantics, Vol1: Fundamentals*. Number 13 in CSLI Lectures Notes Series. Stanford: Center for the Study of Language and Information: University of Chicago Press.
- POLLARD C. & SAG I. (1994). *Head-Driven Phrase Structure Grammar*. University of Chicago Press and CSLI Publications.
- PUSTEJOVSKY J. (1994). Semantic typing and degrees of polymorphism. *Current Issues in Mathematical Linguistics*.
- PUSTEJOVSKY J. (1995). *The Generative Lexicon*. MIT Press.
- PUSTEJOVSKY J. & ANICK P. (1988). On the semantic interpretation of nominals. In *Proceedings of COLING*, Budapest.
- TENNY C. (1994). *Aspectual Roles and the Syntax-semantic Interface*. Dordrecht : Kluwer.
- WEINREICH U. (1964). Webster's third : A critique of its semantics. *International Journal of American Linguistics*, **30**, 405–409.