
Un cas de « polysémie logique » : modélisation de noms d'action en français ambigus entre processus et artefact

Evelyne Jacquy

UMR ATILF - CNRS - Nancy2 - Nancy1 / 44, avenue de la Libération / BP 30687 /
F-54000 NANCY

Evelyne.Jacquy@atilf.fr

RÉSUMÉ. Dans cet article, nous proposons une modélisation logique d'une classe de noms d'action en français qui sont lexicalement ambigus entre une interprétation processive et une interprétation résultative dans laquelle ils désignent des objets créés par le processus sous-jacent, des artefacts. Nous expliquons en quoi ce type de noms d'action relève de la polysémie logique et proposons de modéliser leur contenu sémantique en étendant des travaux antérieurs. Ces travaux, basés sur la notion de type pointé et ses extensions en λ -calcul typé, fournissent un cadre de formalisation satisfaisant une approche compositionnelle du sens et allant dans le sens d'un traitement de l'ambiguïté lexicale en français qui soit à la fois unifié, opératoire et conforme aux propriétés linguistiques observées.

ABSTRACT. This paper describes a modeling in a logical framework of the semantic content of a class of action nouns in French that are ambiguous between a processive interpretation and a resultative one in which these nouns refer to new objects in the world that are created by the process - artefacts. The first part of the paper is about an explanation of the properties that show why these ambiguous action nouns can be considered as logical polysemic words. The second part deals with the necessary extensions that have been realised in order to modelise their ambiguous semantic content. These extensions are added to a previous modelisation based on the notion of dotted types and their modelings in the framework of typed λ -calculus.

MOTS-CLÉS : sémantique lexicale, ambiguïté processus/artefact, polysémie logique, type pointé, λ -calcul typé

KEYWORDS: Lexical semantics, process/artefact ambiguity, logical polysemy, dotted type, typed λ -calculus

1. Introduction

Selon (Kleiber, 1999), l'ambiguïté lexicale peut se définir comme la propriété, pour de nombreux mots dans les langues, de partager des interprétations différentes qui sont observables dans des contextes linguistiques non ambigus.

La modélisation du phénomène de l'ambiguïté lexicale est une problématique du TAL pour plusieurs raisons. Tout d'abord, ce phénomène est productif : un comptage simple¹ sur les 46 094 noms du TLF (CNRS, 1976-1994) montre que, dans près de la moitié des cas (47,24 %), ceux-ci peuvent être considérés comme lexicalement ambigus dans la mesure où ils admettent plus d'une interprétation². Ensuite, le phénomène de l'ambiguïté lexicale est caractéristique des langues et, si l'on s'inscrit dans une perspective où les applications en TAL doivent demander une adaptation minimale de la part des utilisateurs, alors il apparaît nécessaire que ces applications soient capables de traiter l'ambiguïté lexicale, ce qui suppose une modélisation préalable. Enfin, l'ambiguïté lexicale est un phénomène propre aux lexiques des langues et, selon (Pustejovsky, 1995), est la cause de nombre de difficultés dans la réalisation de lexiques de langues qui soient à la fois opératoires et valides du point de vue linguistique³. Or l'absence de ressources lexicales en français à large couverture, ou leurs difficultés d'exploitation dans le cadre du TAL, est une des raisons qui expliquent la difficulté qu'il peut y avoir à construire des applications à large couverture.

Parmi les types d'ambiguïté lexicale, nous fondant sur (Kleiber, 1999) et (Cruse, 1986) ; (Cruse, 1995), nous avons défendu l'idée dans (Jacquey, 2001) que la polysémie logique, illustrée par des noms prototypiques tels que *livre* (objet physique vs. informationnel) ou *fenêtre* (objet physique vs. ouverture) en français, se situe entre l'imprécision et la polysémie au sens strict car elle partage des propriétés de chacune de ces deux classes couvrant une majorité des cas d'ambiguïté lexicale. À partir de ce positionnement intermédiaire de la polysémie logique, on peut émettre l'hypothèse

1. Afin de réaliser ce comptage, nous avons extrait l'ensemble des définitions de noms que l'on trouve dans le TLF, l'un des dictionnaires de référence sur le français du XIX^e et du XX^e siècle. L'extraction a été réalisée à partir de la version XML du dictionnaire en différenciant les définitions qui décrivent les sens des mots vedette et celles qui décrivent les sens de locutions dans lesquelles les mots vedette apparaissent. En mettant de côté les définitions de locutions, cette extraction a permis de constituer un corpus de 125 314 définitions (environ 67 % du total des définitions pour les noms) qui décrivent les sens de 46 094 noms du français. Une répartition par tranches montre ensuite que, parmi les 47,24 % de noms considérés comme lexicalement ambigus, on trouve 21 733 noms à 2 définitions et plus (47,24 %), 13 468 à 3 définitions et plus (29,22 %), 9 231 à 4 définitions et plus (20,03 %), 6 700 à 5 définitions et plus (14,54 %), pour terminer à un substantif, *point* à 82 définitions.

2. Nous faisons ici l'hypothèse que chaque définition indique une interprétation particulière d'un substantif.

3. Le caractère opératoire d'un lexique de langue peut par exemple se mesurer selon sa capacité à prédire les types d'arguments d'un prédicat : on peut *lire un livre* mais pas *lire une maison* et, dans des contextes très particuliers, on peut *lire un mur* (par exemple le mur du métro). La validité linguistique peut se mesurer relativement à la conformité du lexique produit avec les faits de langue observés et au degré de couverture de la langue qu'il reflète.

que sa modélisation offre la perspective d'une description cohérente des mots lexicalement ambigus du français. En effet, à partir d'une modélisation de la polysémie logique, une modélisation de l'imprécision d'une part et de la polysémie au sens strict d'autre part, pourraient être envisagées par le blocage des propriétés non partagées par chacune de ces deux classes respectivement. C'est donc dans cette perspective que nous avons proposé dans (Jacquey, 2004), dans le cadre d'une approche compositionnelle du sens, une modélisation qui était basée d'une part sur la création d'un type complexe dans (Pustejovsky, 1995) (section 3.2.1) destiné à représenter les sens possibles des polysèmes logiques, et d'autre part sur son raffinement dans deux travaux récents, (Pinkal *et al.*, 2000) et (Asher *et al.*, n.d.).

Dans cet article, nous proposons une extension de cette modélisation initiale avec l'objectif d'évaluer l'expressivité de la modélisation initiale de (Jacquey, 2004). L'extension décrite ici s'attache à représenter le contenu sémantique des noms d'un type particulier d'ambiguïté lexicale, l'ambiguïté processus/artefact dont relèvent des noms tels que *chargement, conquête, construction, création, description, inscription, invention, reproduction*, etc. Après avoir montré que le fonctionnement sémantique de ces noms est conforme à celui des polysèmes logiques (section 2), nous détaillons les adaptations de la modélisation initiale qui sont nécessaires pour rendre compte des propriétés spécifiques des noms ambigus entre processus et artefact.

2. L'ambiguïté processus/artefact : un cas de polysémie logique

Les *polysèmes logiques*, dans la terminologie de (Pustejovsky, 1995) ; (Asher *et al.*, n.d.) ; (Pinkal *et al.*, 2000), sont aussi analysés comme des *mots à facettes sémantiques* dans (Cruse, 1986) ; (Cruse, 1995) et (Copestake *et al.*, 1995) ou encore comme des *noms multitypés* dans (Godard *et al.*, 1996). Ils sont aussi considérés comme des cas particuliers de métonymie dans (Kleiber, 1999) et dans la tradition lexicographique. Au-delà des divergences terminologiques, ces auteurs s'accordent sur le fait que les sens des mots de cette classe partagent trois propriétés caractéristiques :

– **la distinction** : les différentes interprétations sont plutôt associées à des sens différents qu'à des cas de spécification sémantique ;

– **la coopération** : dans ce que nous appellerons par la suite une *coprédication*, plusieurs prédicats sémantiquement non ambigus semblent pouvoir sélectionner la même occurrence, ou plusieurs occurrences référentiellement liées, d'un polysème logique⁴ ;

4. Comme nous y invite l'un des rapporteurs de l'article dans la revue, nous indiquons dès maintenant que cette propriété est variable selon les substantifs observés. Si cette propriété est vraie pour la majeure partie des emplois du mot *livre* par exemple, son degré d'acceptabilité est moins évident avec certains emplois du mot *journal* dans un emploi tel que (??) *Ce journal a massivement embauché depuis 1946 et est très abîmé* alors que chacune des propositions coordonnées est tout à fait acceptable *Ce journal a massivement embauché depuis 1946, Ce journal est très abîmé*. Sur cet exemple néanmoins, on pourrait objecter que les sens que l'on

– **la portée sémantiquement variable de la quantification** : dans une coprédication quantifiée, le quantificateur porte préférentiellement sur l'un des sens possibles des polysèmes logiques.

Cependant, ces trois propriétés ont en grande partie été illustrées pour le nom prototypique *livre*. De plus, la plupart des travaux traitant de cette question portent sur l'anglais. Afin de pouvoir étudier ce phénomène sur le français, nous avons interrogé l'une des ressources de référence sur cette langue, le TLF.

Parmi les classes identifiées par Pustejovsky comme relevant de la polysémie logique, nous avons sélectionné celle de l'ambiguïté processus/résultat et au sein de celle-ci, nous nous sommes concentrés sur les cas d'ambiguïté processus/artefact (le résultat dénote l'objet artefactuel produit par le déroulement du processus) et cela pour plusieurs raisons. Premièrement, à l'inverse de l'image conceptuelle homogène que l'on peut se faire de l'entité à laquelle réfère le nom *livre* selon (Kleiber, 1999) et (Cruse, 1995), les entités auxquelles réfèreraient un nom ambigu entre processus et artefact auraient clairement une image conceptuelle bipartite avec une facette relevant des processus et une facette relevant des objets. Deuxièmement, ce type de noms est intéressant car, comme nous le verrons en (2.2) et (2.3), les sens de processus et d'artefact semblent partager les propriétés des polysèmes logiques prototypiques comme *livre*, mais ne le font pas à l'identique. Troisièmement, les noms ambigus entre processus et – résultat ou artefact – sont, pour une grande partie d'entre eux, morphologiquement liés à un verbe, que ce soit par suffixation ou par conversion. Enfin, cette classe est une de celles qui sont le plus facilement identifiables d'un point de vue méthodologique dans la ressource qui nous a servi de référence sur le français. Toujours sur le plan méthodologique, la différence entre contexte processif et contexte artefactuel est plus aisée à reconnaître que celle entre contexte processif et contexte statif.

Dans la suite, nous examinerons la question du classement de ce type d'ambiguïté lexicale en tant que cas de polysémie logique après en avoir donné quelques exemples en français.

2.1. *L'ambiguïté processus/artefact en français*

En première approche, l'ambiguïté processus/résultat est repérable dans le dictionnaire TLF, plus précisément dans le corpus constitué par l'ensemble des définitions de mots vedette dans le TLF, en sélectionnant les noms dont une définition contient à la fois la suite *Action de V_inf*, le verbe à l'infinitif étant morphologiquement lié

cherche à faire collaborer sont d'organisation humaine et celui d'exemplaire physique contenant de l'information produite par l'organisation. Cette ambiguïté, organisation humaine/exemplaire physique, se retrouve notamment avec des exemples portant sur le nom propre *Honda* dans (Pustejovsky, 1995) mais les emplois cités ne font justement pas intervenir la coprédication. Disons pour conclure que la propriété de coprédication varie non seulement en fonction des substantifs, mais aussi en fonction des sens que l'on cherche à faire coopérer.

au substantif, et la suite *résultat de cette action*⁵. Selon ce critère, le TLF compte 1 098 noms qui sont de bons candidats pour figurer parmi les cas d'ambiguïté processus/résultat. Parmi ces noms ambigus, la recherche des noms relevant de l'ambiguïté processus/artefact a été réalisée en utilisant la présence dans les définitions de la suite *Ce qui est V_part.passé*, le verbe étant morphologiquement lié au substantif⁶. Cette seconde recherche a conduit à la sélection de 58 noms potentiellement ambigus entre un processus et l'objet artefactuel produit par ce processus.

La faible productivité apparente du phénomène de l'ambiguïté processus/artefact demande à être commentée. En fait, cette petite quantité de noms relevant potentiellement de ce type d'ambiguïté est due au faible degré de rappel des critères, et plus particulièrement de celui qui vise l'interprétation artefactuelle des substantifs. Un examen manuel des définitions des noms ambigus entre processus et résultat montre que l'artefact est identifiable par d'autres suites de caractères comme par exemple *chose, contrat, objet, ouvrage, représentation*, etc. Cependant, certains termes présents dans les définitions étant eux-même ambigus (par exemple *représentation*), nous avons préféré éviter au maximum le bruit. En tout état de cause, la délimitation précise du phénomène de l'ambiguïté processus/artefact en français demande encore des recherches dans le cadre cette fois de la classification et de l'acquisition de l'information étant donnée l'hétérogénéité des indications que l'on peut trouver dans la ressource de référence utilisée⁷.

Enfin, pour compléter ce commentaire, rappelons que l'ambiguïté processus/artefact est une sous-classe de la classe processus/résultat, elle-même comptant comme une classe parmi les sept identifiées par Pustejovsky⁸. À la suite de Kleiber et de Pustejovsky qui considère ces classes comme des métonymies, une recherche parmi les noms du TLF montre qu'environ 6 000 d'entre eux ont des emplois métonymiques, soit 12,4 % de l'ensemble, ce qui en fait une partie non négligeable du vocabulaire du français. Or, si la modélisation proposée dans (Jacquey, 2004) peut être étendue pour rendre compte de l'ambiguïté processus/artefact, et donc processus/résultat, les extensions propres aux autres classes de polysémie logique sont peut-être envisageables.

5. Le critère décrit ici est conditionné par la façon dont le sens processif des noms d'action est décrit dans le dictionnaire. Il n'est donc pas revendiqué comme universel mais simplement adapté à la ressource sur laquelle nous nous appuyons.

6. On rappelle à nouveau que ce critère est lui aussi fortement dépendant du mode de rédaction des définitions dans le TLF.

7. Une telle hétérogénéité n'est d'ailleurs pas surprenante quand on sait que la rédaction du TLF s'est étalée sur une trentaine d'années et a mis à contribution une centaine de rédacteurs environ. Bien que cet état de fait soulève la question de l'exploitabilité des dictionnaires de langue, il ne conduit pas pour autant à une réponse négative, en particulier si l'on prend en compte la couverture et la richesse offertes par les dictionnaires.

8. En dehors de l'ambiguïté processus/résultat, Pustejovsky considère comme relevant de la polysémie logique les ambiguïtés massif/comptable, contenant/contenu, figure/fond, lieu/collectif_humain, plante/nourriture, produit/producteur.

2.2. Propriétés communes aux polysèmes logiques et aux noms d'action ambigus processus/artefact

Afin de faire apparaître plus simplement les propriétés de deux noms d'action ambigus entre processus et artefact, *construction* et *reproduction*, différents exemples construits ont été produits à partir d'exemples attestés reproduits ci-dessous.

- (1) a. La maison fondée par M. Olier, en 1645, n'était pas la grande **construction** quadrangulaire, à l'aspect de caserne, qui forme maintenant un côté de la place Saint-Sulpice. Renan, Souvenirs d'enfance et de jeunesse, 1883, p. 265.
- b. Il serait intéressant de vérifier si les lignes régulières ne sont que dans le cerveau de l'homme. Les animaux ne les reproduisent pas dans leurs **constructions**, ou plutôt dans les ébauches de régularité que présentent leurs ouvrages, comme le cocon, l'alvéole. E. Delacroix, Journal, 1863, p. 428.
- c. La **construction** du nouvel édifice fut confiée à l'architecte Rossellini R. Ménard, Hist. des B.-A., 1882, p. 150.
- (2) a. Elle habitait, dans un hôtel de la rue Saint-Sulpice, une petite chambre très bleue ; elle avait accroché aux murs des **reproductions** de Cézanne, de Renoir, du Gréco, et les dessins d'un ami espagnol qui voulait peindre (Beauvoir, Mém. j. fille, 1958, p. 284).
- b. Les rédacteurs sont chargés d'assurer la **reproduction** des séances publiques, conformément aux règles prescrites par l'Assemblée. Ils surveillent les copies, les envoient à l'imprimeur, corrigent les épreuves ; ils prennent les mêmes soins pour les procès-verbaux des séances secrètes, quand l'Assemblée en ordonne l'impression (Règlement Ass. nat., 1849, p. 41).

Les noms d'action ambigus semblent partager les trois propriétés distinctives des polysèmes logiques rappelées ci-dessus. Tout d'abord, les deux interprétations qu'ils admettent relèvent de sens différents qui apparaissent dans des contextes identifiables. L'interprétation processive est caractérisable grâce à des indicateurs aspectuels⁹. L'interprétation processive de ces noms d'action peut toujours apparaître avec des prédicats verbaux typiques comme *avoir lieu + date*, *durer*, *commencer*, *effectuer*, *procéder à*, etc.

- (3) a. Les rédacteurs ont procédé à la reproduction_[proc] des séances publiques.
- b. La construction_[proc], confiée à Rossellini, a débuté à la fin du XIX^e siècle.

La caractérisation de l'interprétation résultative est plus complexe car plus diversifiée mais elle est, d'une manière générale, reflétée par des éléments contextuels désignant des propriétés des objets construits par le processus. Dans l'exemple suivant,

9. On peut se reporter à (Vendler, 1967) ; (Verkuyl, 1989) ; (Godard *et al.*, 1996) pour une description de ces indicateurs et de leur fonctionnement.

l'interprétation résultative de *reproduction* est repérable grâce au prédicat de localisation *être accroché aux murs* et celle de *construction* l'est grâce au prédicat *former* + *NP_lieu* qui désigne une propriété du bâtiment d'habitation résultant.

- (4) a. [...] elle avait accroché aux murs des reproductions_[art] de Cézanne, de Renoir, du Gréco, [...]
 b. La maison de M. Olier n'était pas la grande construction_[art] quadrangulaire qui forme un côté de la place Saint-Sulpice.

Ensuite, ces noms d'action admettent la coopération de leurs sens.

- (5) a. Les reproductions_[proc] des séances publiques sont effectuées conformément aux règles prescrites par l'Assemblée, puis _[art] directement envoyées aux imprimeurs.
 b. Il serait intéressant de vérifier si les animaux ne reproduisent pas des régularités humaines dans leurs constructions_[proc et/ou art ?] qui _[art] sont par exemple le cocon, l'alvéole [...]
 c. Il serait intéressant de vérifier si les animaux ne reproduisent pas des régularités humaines dans les constructions_[proc] qu'ils effectuent annuellement parmi lesquelles_[art et/ou proc ?] on compte le cocon, l'alvéole [...]

En (5), le groupe nominal *les reproductions des séances publiques* est sous l'influence de deux prédicats, *être effectué* et *être envoyé*. Chacun de ces prédicats portant respectivement sur un processus et sur un objet physique, le groupe *les reproductions des séances publiques* est donc interprété sous ces deux angles simultanément. L'analyse de l'exemple (5b) sera plus nuancée. Dans la première partie de la phrase, le groupe *leurs constructions* peut être interprété comme un processus, mais aussi comme un artefact (*construction* serait ici remplaçable par *leur habitat*). En (5c), la surspécification du processus au moyen de la relative *qu'ils effectuent annuellement* ne bloque pas la coprédication mais celle-ci est plus acceptable si on la modifie en remplaçant l'anaphorique *qui* par *parmi lesquelles* suivi de *on compte*. La coopération des sens de processus et d'artefact est donc à considérer ici comme une possibilité.

Enfin, lorsque les noms d'action ambigus entre processus et artefact apparaissent dans des coprédications quantifiées, le sens sélectionné par le prédicat principal se trouve préférentiellement sous la portée du quantificateur universel.

- (6) a. Toutes les reproductions de séances publiques qui sont effectuées selon les règles prescrites par l'Assemblée sont directement envoyées à l'imprimeur.
 b. Toutes les constructions qui sont effectuées par les animaux reproduisent des régularités humaines.

Dans une interprétation possible de la phrase (6a), la quantification universelle, réalisée par *toutes les*, porte sur le sens artefactuel de *reproduction* et une quantification existentielle, sur le sens processif. Une glose possible de (6a) serait « *Pour tout*

artefact envoyé à l'imprimeur, il existe un processus de reproduction dont l'artefact est un résultat et qui suit les règles prescrites par l'Assemblée ».

2.3. Différences entre polysèmes logiques et noms d'action ambigus processus vs. artefact

Le parallélisme entre les propriétés des polysèmes logiques et celles des noms d'action n'est pas total. Deux différences au moins peuvent être soulignées.

La propriété de la coprédication n'est pas aussi facilement satisfiable par les noms d'action ambigus processus vs. artefact. Alors que les trois configurations linguistiques observées par Jayez et Godard – la coordination, la reprise anaphorique et l'insertion d'une relative – sont parfaitement admises avec *livre*, on constate des degrés d'acceptabilité moindres avec les noms d'action ambigus, notamment avec la coordination où, de plus, la détermination semble jouer un rôle non négligeable.

- (7) a. ? Les reproductions de Cézanne sont accrochées au mur et ont été effectuées il y a peu.
b. Ces reproductions de Cézanne sont accrochées au mur et ont été effectuées il y a peu.
c. ?? Des reproductions de Cézanne sont accrochées au mur et ont été effectuées il y a peu.
- (8) a. ? Les constructions des animaux sont effectuées chaque année et reproduisent des régularités humaines.
b. Ces constructions des animaux sont effectuées chaque année et reproduisent des régularités humaines.
c. ? Des constructions des animaux sont effectuées chaque année et reproduisent des régularités humaines.

Comme nous l'avons suggéré dans la section précédente, les acceptabilités des coprédications faisant coopérer le sens processif et le sens artefactuel sont possibles mais pas systématiques.

Ensuite, l'ambiguïté processus/artefact pour les noms d'action pris en compte n'est pas présente dans toutes leurs occurrences. Une étude de la corrélation entre leurs propriétés argumentales et aspectuelles¹⁰ a montré qu'ils étaient effectivement ambigus essentiellement dans deux types de configurations.

10. Cette étude a été faite à partir des travaux de plusieurs auteurs, (Dowty, 1982), (Vendler, 1967) et (Verkuyl, 1989) pour les propriétés aspectuelles et (Milner, 1982), (Samvelian, 1995) pour les propriétés argumentales, et (Grimshaw, 1990) pour les propriétés aspectuelles et argumentales.

– S'ils désignent un processus de création d'un artefact, dérivé à partir d'un objet préexistant, ils sont ambigus dès que leur argument agentif n'est plus réalisé. Ainsi, le groupe nominal *les reproductions des séances publiques* est ambigu alors que *les reproductions des séances publiques par les rédacteurs* est plutôt interprété comme processif.

– S'ils désignent un processus de création d'un artefact ne pouvant pas être considéré comme dérivé d'un objet préexistant, seule leur construction absolue (aucun argument réalisé) est ambiguë. Les groupes nominaux *la construction de la maison par les ouvriers* et *la construction de la maison* ne sont pas compatibles avec des contextes résultatifs tels que *est solide*, *est quadrangulaire*, alors que le groupe *la construction* est compatible avec des contextes aussi bien processifs que résultatifs.

2.4. Conclusion sur le classement des noms d'action ambigus processus vs. artefact

Pour conclure sur cette question, les noms d'action ambigus entre processus et artefact nous semblent devoir être considérés comme des polysèmes logiques car ils semblent davantage partager leurs trois propriétés caractéristiques qu'ils ne semblent s'en écarter totalement. Comme cela a été montré dans les exemples précédents, l'interprétation processive et l'interprétation résultative proviennent de sens distincts, observables dans des contextes non ambigus. Ces deux interprétations peuvent coopérer au sein de coprédications même si cette seconde propriété est acceptée dans une moindre mesure. Enfin, dans une coprédication quantifiée, le sens sélectionné par le prédicat principal semble être préférentiellement sous le champ du quantificateur universel plutôt que le sens complexe, processus et artefact à la fois.

Nous l'avons vu cependant, le parallélisme entre polysèmes logiques prototypiques et noms d'action ambigus entre processus et artefact n'est pas total. Cet état de fait apporte un second argument au rapprochement. Comme nous le verrons dans la section (4.1), l'une des deux différences soulignées – la limitation de cette ambiguïté à certaines configurations argumentales – peut être modélisée par le biais d'une extension spécifique de la modélisation initiale de la polysémie logique telle que nous l'avons reprise de travaux antérieurs (Jacquey, 2004).

Enfin, pour terminer sur ce point, nous rappellerons que l'un de nos objectifs dans la modélisation de l'ambiguïté lexicale sémantique en français est d'avancer des pistes de solutions pour la réalisation de lexiques de langue valides et opératoires. Pour cela, il nous semble plus pertinent de rapprocher les noms ambigus entre processus et artefact des polysèmes logiques. En effet, compte tenu de la diversité des emplois et du polymorphisme lexical, et avec le souci de parvenir à faire émerger des axes de classification globaux, quoique certainement parfois trop imprécis, pour les 12,4 % de noms du TLF que nos heuristiques (utilisation d'indicateurs d'emplois comme *par métonymie* en première approche) retiennent comme candidats au statut de polysème logique, il semble nécessaire de faire émerger un mécanisme sémantique commun afin de mieux faire apparaître les différences. Enfin, il est aussi possible de voir la question

sous la forme d'une échelle, les polysèmes prototypiques comme *livre* affichant toutes les propriétés parfaitement satisfaites, les noms ambigus entre processus et artefact les satisfaisant de manière moins évidente, mais relevant néanmoins d'un mécanisme sémantique similaire.

3. Modélisation de la polysémie logique

Nous nous tournons maintenant vers les aspects plus formels de cet article en rappelant les éléments fondamentaux de la modélisation de la polysémie logique sur laquelle nous nous appuyons pour représenter le contenu sémantique particulier des noms d'action ambigus entre processus et artefact. Nous rappelons brièvement le cadre général formel dans lequel s'inscrit cette modélisation (section 3.1). Nous poursuivons avec un rappel des principales extensions au cadre formel qui nous ont permis de modéliser le fonctionnement sémantique des polysèmes logiques dans un cadre unifié, linguistiquement valide et plus opératoire (section 3.2).

3.1. Cadre formel : λ -calcul typé étendu à la quantification et sémantique lexicale

À partir des livres de référence, (Hindley *et al.*, 1986) et (Dowty *et al.*, 1985), nous avons extrait les points qui nous semblent indispensables à la suite de l'exposé.

Dans le cadre de la sémantique de Montague, chaque mot ou syntagme d'un énoncé est représenté par un terme. Les termes du λ -calcul sont de quatre formes combinables entre elles : les *variables*, notées x, y, z, u, v pour les *variables élémentaires* ou P, Q, R pour les *variables fonctionnelles*, les *constantes*, notées a, b, c, d , les *applications*, notées (MN) et les *abstractions*, notées $\lambda x.M$. L'application (MN) est équivalente au calcul d'une fonction M en un point N et l'abstraction $\lambda x.M$ est une notation mathématique permettant d'écrire toutes les fonctions de manière générique.

De plus, chaque terme est associé à un type qui dépend de la forme du terme et de la catégorie syntaxique de l'item qu'il représente, ces associations de types étant mémorisées dans l'ensemble A qui représente le contexte d'assignation de types. De même que les termes, les types sont construits à partir de types atomiques, le type e des entités et le type t des valeurs de vérité. Le constructeur \rightarrow permet de construire les types plus complexes. Le typage des termes dépendant de leur structure interne, quatre règles d'inférence permettent de construire le type des termes, en particulier celui des abstractions et des applications :

$$\begin{array}{ll}
 \text{constantes} & \frac{[c:\alpha] \in A}{A \vdash c:\alpha} \\
 \text{application} & \frac{A \vdash M:\alpha \rightarrow \gamma \quad A \vdash N:\alpha}{A \vdash (MN):\gamma} \\
 \text{variables} & \frac{[x:\alpha] \in A}{A \vdash x:\alpha} \\
 \text{abstraction} & \frac{A \vdash x:\alpha \quad A \vdash M:\gamma}{A \vdash \lambda x.M:\alpha \rightarrow \gamma}
 \end{array}$$

[1]

Dans ces règles, les symboles α et γ représentent des variables de type, c'est-à-dire des types non encore connus et qui seront remplacés par les contraintes de types fournies par les constantes du lexique et les équations de types issues de l'application de ces règles. La règle de l'abstraction construit le type d'une λ -abstraction $\lambda x.M$ en fonction du type de la variable x et de celui du terme M . Cette règle peut être représentée par l'équation de types $type(\lambda x.M) = type(x) \rightarrow type(M)$. À l'inverse, la règle de l'application, entre un terme M , de type fonctionnel $\alpha \rightarrow \gamma$, et un terme N , de type α , construit le type γ pour le terme (MN) si les deux occurrences de la variable de type α reçoivent des contraintes de types qui sont *identiques ou unifiables*. Si ce n'est pas le cas, on dit qu'il y a *conflit de types* et c'est dans ce cas qu'il faut définir une procédure de résolution.

Enfin, pour symboliser les restrictions sélectionnelles des prédicats, le type e est subdivisé en de nombreux sous-types organisés hiérarchiquement. Ainsi, le prédicat *être effectué*, assimilé ici au fonctionnement d'un adjectif, porte préférentiellement sur des variables de type *processus*, type abrégé par *proc*. Dans une phrase comme (5a), il serait représenté par un terme de la forme $\lambda R \lambda v. effectué(v) \wedge R(v)$. À partir des règles d'abstraction et d'application, on déduit que ce terme est de type $(proc \rightarrow t) \rightarrow proc \rightarrow t$.

Les mêmes règles de typage permettent d'établir un ensemble d'équations de types pour une phrase complète. Ainsi, dans la phrase *Les rédacteurs ont envoyé les reproductions effectuées*, l'ensemble des équations de types pourrait être le suivant :

$$\begin{aligned} type(effectué) &= type(reproduction) \rightarrow type(effectué reproduction) \\ type(le) &= type(effectué reproduction) \rightarrow type(le(effectué reproduction)) \\ type(le(effectué reproduction)) &= \\ type(avoir envoyé) &\rightarrow type((le(effectué reproduction)) avoir envoyé) \\ type((le(effectué reproduction)) avoir envoyé) &= type(les rédacteurs) \rightarrow \\ type(((le(effectué reproduction)) avoir envoyé) les rédacteurs) & \end{aligned}$$

À partir de ces équations et à partir des types fournis par les constantes, il est possible d'estimer le nombre et la nature des conflits de types.

3.2. Éléments fondamentaux de la modélisation

Avant d'entrer dans le détail de ces éléments, nous revenons sur la manière dont les propriétés des polysèmes logiques ont influé sur la conception de la modélisation initiale (Jacquey, 2004) ainsi que sur les travaux sur lesquels elle se fonde (Pustejovsky, 1995), (Asher *et al.*, n.d.) (AP) et (Pinkal *et al.*, 2000) (PK).

Les deux premières propriétés – distinction et coopération possible des sens des polysèmes logiques – ont conduit d'une part à la définition de sens séparés, mais d'autre part, à l'exclusion de toute modélisation reposant sur la disjonction entre entrées lexicales ou bien entre acceptions à l'intérieur d'une même entrée. La coopéra-

tion possible des sens impose en effet de pouvoir accéder, non seulement à chacun d’eux, mais aussi à l’ensemble formé par ces sens, à tout moment de l’analyse de l’énoncé.

Afin de rendre compte de ces deux besoins apparemment contradictoires – distinction des sens et disposition simultanée et permanente de ceux-ci – la notion de type pointé, telle qu’elle est définie par (Pustejovsky, 1995) et raffinée par (AP) et (PK), nous a semblé la plus adéquate¹¹. Ce nouveau type complexe constitue un moyen d’accéder, à partir d’un même point d’entrée, le type pointé, à la fois aux sens simples d’un polysème logique et à son sens complexe.

La dernière propriété – les interprétations variables des coprédications quantifiées – a conduit à privilégier un traitement garantissant l’accès à chacun des sens possibles des polysèmes, indépendamment de l’accès à l’ensemble de ces sens, c’est-à-dire au type pointé. En effet, pour que la quantification puisse porter alternativement sur l’un ou l’autre des sens simples d’un polysème logique, il est nécessaire que le traitement permette l’accès aux variables de type simple indépendamment de l’accès à la variable de type pointé.

Les sections suivantes (3.2.1, 3.2.2, 3.2.3)¹² ont pour but de rappeler les travaux antérieurs à la modélisation (Jacquey, 2004) avant d’aborder les extensions faites à cette dernière afin de rendre compte des propriétés particulières des noms ambigus entre processus et artefact.

Afin de montrer l’ensemble des aspects à représenter, ces rappels porteront sur l’ambiguïté lexicale du nom prototypique *livre* puisque celui-ci satisfait toutes les propriétés des polysèmes logiques. La phrase étalon à laquelle chacun des travaux cités est comparée est la coprédication *Marie a traduit le livre rouge*. Dans cette phrase, la même occurrence du nom *livre* est respectivement sous l’influence du prédicat *traduire*, qui sélectionne un objet informationnel, et celle du prédicat *rouge* qui modifie plutôt un objet *physique*. Certains des aspects à représenter pourraient se révéler non nécessaires avec les noms ambigus entre processus et artefact, en particulier la nécessité de rendre compte de la portée sémantiquement différente de la quantification puisque, dans tous les cas, un seul processus a produit un seul artefact. Le problème du nombre d’entités désignées selon le sens sélectionné avec *livre*, unique avec le sens informationnel, indéterminé avec le sens physique, ne se pose donc pas de la

11. En particulier, une approche par disjonction des sens avec unification des traits caractéristiques de chacun d’eux interdirait la représentation des coprédications. D’un autre côté, une approche bâtie sur l’unification du signifiant du nom ambigu, sans imposer dans le même temps la compatibilité des traits caractéristiques de chaque sens, permettrait de produire une représentation bien formée d’une phrase très difficilement acceptable comme ? ? *la construction par les ouvriers forme l’angle de la place Saint-Sulpice*.

12. Nous précisons sur ce point particulier que l’exposé des travaux (Pustejovsky, 1995, Pinkal *et al.*, 2000, Asher *et al.*, n.d.) que l’on peut lire dans cet article a déjà été publié dans (Jacquey, 2004). Les améliorations qui apparaissent sont dues à la relecture attentive et rigoureuse des correcteurs de la revue, travail fastidieux mais très utile pour lequel nous les remercions chaleureusement.

même manière avec l'ambiguïté processus/artefact. Cependant, toujours dans la perspective d'appliquer un mécanisme sémantique commun à l'ensemble des polysèmes logiques, mais aussi parce que la prise en compte de cette propriété, semble-t-il moins nécessaire avec l'ambiguïté processus/artefact, ne conduit pas à des résultats linguistiquement faux, cet aspect est conservé.

3.2.1. *Le lexique génératif : une première approche de la notion de type pointé*

La notion de type pointé a été initialement définie par Pustejovsky dans (Pustejovsky *et al.*, 1988), (Pustejovsky, 1991), (Pustejovsky, 1995) et (Pustejovsky, 1996). Selon lui, un type pointé est associé à un mot ambigu lorsque celui-ci appartient à la catégorie des polysèmes logiques. Leurs sens sont toujours accessibles cognitivement et admettent la coprédication.

Le fait d'associer un type pointé $\sigma_1 \bullet \sigma_2$ aux polysèmes logiques revient selon (Pustejovsky *et al.*, 1988) à considérer qu'ils dénotent non seulement chacun des sens symbolisés par les types simples, σ_1 et σ_2 , mais aussi un troisième sens qui correspond à la combinaison des deux premiers et qui est symbolisé par le type pointé $\sigma_1 \bullet \sigma_2$. L'ensemble de sens possibles est considéré comme un paradigme, appelé un *paradigme lexical conceptuel* et noté $\sigma_1 \bullet \sigma_2_LCP$ (*Lexical Conceptual Paradigm*). Pustejovsky introduit un nouveau constructeur qu'il appelle du même nom, *lcp*. Ce constructeur est introduit par le biais de la règle suivante :

$$\frac{M : \sigma_1 \quad M : \sigma_2}{lcp(M) : \sigma_1 \bullet \sigma_2}$$

Comme le montre cette nouvelle règle, si un terme M est de type σ_1 et de type σ_2 , l'application du constructeur *lcp* au terme M permet de lui associer le type $\sigma_1 \bullet \sigma_2$. Pour obtenir ensuite l'ensemble du paradigme lexical conceptuel, Pustejovsky définit la règle d'inférence *Type Pumping*, (Pustejovsky, 1995), p. 120. Cette règle est contrôlée par un opérateur Σ nommé *Type Pumping Operator* qui est dédoublé. L'opérateur Σ_1 , correspondant à Σ restreint au type σ_1 , extrait le type simple σ_1 à partir du type pointé $\sigma_1 \bullet \sigma_2$. De même, l'opérateur Σ_2 extrait le type simple σ_2 à partir du type pointé. L'application de cette règle permet donc d'écrire le faisceau de types tel que :

$$\sigma_1 \bullet \sigma_2_lcp = \{\sigma_1 \bullet \sigma_2, \Sigma_1[\sigma_1 \bullet \sigma_2], \Sigma_2[\sigma_1 \bullet \sigma_2]\} = \{\sigma_1 \bullet \sigma_2, \sigma_1, \sigma_2\}$$

Le prédicat *livre*, étant considéré comme un polysème logique dans le lexique génératif, est associé au type fonctionnel, $p \bullet i \rightarrow t$, p symbolisant le sens physique et i le sens informationnel. De plus, comme le montre la traduction logique ci-dessous, tous les types pointés sont définis par une relation précise, notée ici R , qui lie les deux variables de type simple, x et y ici. Dans (Pustejovsky, 1995), R est considéré comme une relation de contenance. Dans le lexique génératif, la représentation associée au nom *livre* est un terme de la forme suivante :

$$\lambda x \bullet y. \exists R [livre(x \bullet y) \wedge R(x, y) \dots]$$

et de type $phys \bullet info \rightarrow t$, x étant de type $phys$ et y étant de type $info$. Cependant, (Pustejovsky, 1995) le remarque lui-même (page 149), la définition formelle des types pointés n'est pas suffisante. Pour définir l'abstraction sur des variables pointées telles que $x \bullet y^{13}$, il faut étendre le λ -calcul simplement typé.

Deux travaux postérieurs (Asher *et al.*, n.d., Pinkal *et al.*, 2000) sont venus raffiner et préciser la notion de type pointé, notamment du point de vue de sa définition logique.

3.2.2. Pinkal et Kohlhase (PK) : une vision fonctionnelle du type pointé

(PK) introduisent une logique de traits avec héritage, notée \mathcal{F}^{\leq} , dans laquelle les restrictions sélectionnelles des prédicats logiques sont représentées par des types complexes appelés des types *record*, appellation que nous traduirons par *trait*. Ces types sont construits sur des ensembles d'attributs dont les valeurs sont des entités, type e . Dans ces types *trait*, chaque attribut représente un sens. Par conséquent, les types *trait* des mots lexicalement ambigus comportent plusieurs attributs.

Dans ce cadre, le nom *livre* est associé à un type *trait* à deux attributs, le premier PHYS et le second INFO. Il reçoit deux descriptions équivalentes du point de vue de leur type, mais différentes du point de vue de leur forme. Au *niveau linguistique*, *livre* est représenté par le prédicat *livre'* qui correspond au terme $\lambda u.livre'(u)$ et qui est de type $\{\{PHYS : e, INFO : e\}\} \rightarrow t$. Associée à cette description linguistique (PK) définissent une *description conceptuelle* symbolisée par le terme $\lambda u.livre^*(u.PHYS, u.INFO)$. Le prédicat *livre** rend compte du fait qu'un item à type pointé correspond à un concept complexe à plusieurs facettes ou parties.

La correspondance entre ces deux représentations de *livre* est assurée par un ensemble de règles d'inférences dédiées à la bonne formation et au typage des termes de type *trait*. La clé de voûte de cette correspondance repose sur les deux règles suivantes¹⁴ :

$$\frac{A \vdash M_1 : \alpha_1 \quad \dots \quad A \vdash M_n : \alpha_n}{A \vdash \{\{l_1 = M_1, \dots, l_n = M_n\}\} : \{\{l_1 : \alpha_1, \dots, l_n : \alpha_n\}\}}$$

$$\frac{A \vdash M : \{\{\dots l : \alpha \dots\}\}}{A \vdash M.l : \alpha}$$

La première règle construit un terme associé au type *trait* correspondant. La seconde règle montre comment extraire les parties (ou facettes) d'un terme de ce type. Dans la représentation linguistique de *livre* (le terme $\lambda u.livre'(u)$), la variable u est un terme de type *trait* $\{\{PHYS : e, INFO : e\}\}$. Comme le décrit la seconde règle, les termes $u.PHYS$ et $u.INFO$ sont des sous-termes de u et correspondent respectivement aux facettes physique et informationnelle des entités représentées par u .

13. Pustejovsky assimile les notions de type pointé $\sigma_1 \bullet \sigma_2$ et de variable pointée $x \bullet y$ lorsque x est de type σ_1 et y de type σ_2 .

14. La notation $\{\{l : \alpha\}\}$ représente le type d'un terme de type *trait*, c'est-à-dire d'un terme dans lequel on indique que la facette l est de type α .

Pour résoudre le conflit entre *livre* et *rouge* dans la phrase *Marie a traduit le livre rouge*, (PK) utilisent d'abord cette correspondance entre niveau linguistique et conceptuel. Avec $v : \{\{\text{PHYS} : e\}\}$ et $u : \{\{\text{PHYS} : e, \text{INFO} : e\}\}$, le syntagme *livre rouge* correspond aux termes suivants :

$$\begin{aligned} & (\lambda R \lambda v. R(v) \wedge \text{rouge}'(v))(\lambda u. \text{livre}'(u)) \\ & (\lambda R \lambda v. R(v.\text{PHYS}) \wedge \text{rouge}^*(v.\text{PHYS}))(\lambda u. \text{livre}^*(u.\text{PHYS}, u.\text{INFO})) \end{aligned}$$

La résolution proprement dite des conflits de types passe par l'utilisation de l'unification sur des types *trait* (plus un type *trait* comporte d'attributs, plus il est spécifique) et du sous-typage entre types fonctionnels. Conformément à ces deux dernières règles, le type fonctionnel de *livre* est plus général que celui de *rouge*. La résolution du conflit permet la β -réduction et conduit au terme $\lambda v. \text{rouge}^*(v.\text{PHYS}) \wedge \text{livre}^*(v.\text{PHYS}, v.\text{INFO})$ avec la variable v de type $\{\{\text{PHYS} : e, \text{INFO} : e\}\}$. La résolution du second conflit de la phrase entre *le livre rouge* et le verbe *traduit* suit la même procédure. On obtient donc, pour la phrase *Marie a traduit le livre rouge*, la formule $\exists x (\text{livre}^*(x.\text{PHYS}, x.\text{INFO}) \wedge \text{rouge}^*(x.\text{PHYS}) \wedge \text{traduit}^*(m, x.\text{INFO}))$. Cette formule est de type t , la constante m , représentant l'individu *Marie*, est de type $\{\{\text{HUM} : e\}\}$ et la variable x est de type $\{\{\text{PHYS} : e, \text{INFO} : e\}\}$.

Ce traitement permet de décrire les accommodations de types, c'est-à-dire l'accès aux différents sens de *livre*, dans le cadre des coprédications. (PK) développent en fait une *vision fonctionnelle des types pointés*. Comme le montre la règle d'extraction des sous-termes ci-dessus, un terme de type *trait* correspond, dans le niveau conceptuel, à autant de sous-termes de type simple qu'il y a d'attributs, donc de sens. Chacun de ces sous-termes étant de type e , on peut voir l'extraction des sous-termes comme un *ensemble de fonctions* ayant respectivement pour image l'ensemble d'entités désignées par les sous-termes du terme v de type *trait*. Dans le cas de *livre*, ces fonctions projettent respectivement les parties physique et informationnelle de chaque individu « livre ». Le traitement développé par (PK) est à la fois expressif et robuste.

Cependant, leur vision du type pointé rencontre des difficultés lorsqu'elle est confrontée à la représentation des variations de quantification, comme avec la phrase *Marie a traduit tous les livres rouges*. En effet, cette phrase correspondrait au terme $\forall x (\text{livre}^*(x.\text{PHYS}, x.\text{INFO}) \wedge \text{rouge}^*(x.\text{PHYS}) \rightarrow \text{traduit}^*(m, x.\text{INFO}))$. Or, selon cette formule, le nombre de traductions correspond au nombre d'éléments x , éléments qui sont tous des livres avec une partie physique et une partie informationnelle propre. Ainsi, Marie pourra très bien avoir traduit plusieurs fois la même œuvre. Il est vrai que cela ne rend pas l'interprétation fautive, néanmoins, il serait intéressant de pouvoir déduire l'interprétation la plus pertinente hors contexte qui fait porter la quantification universelle sur le sens informationnel uniquement. Pour ce faire, (PK) proposent de contextualiser les quantificateurs. Cette contextualisation est représentée par une ou plusieurs variables contextuelles. Nous ne discuterons cependant pas l'aspect technique de cette contextualisation. Les variables contextuelles sont estimées sur la base d'un contexte, non précisé, et sur la base d'un ensemble de restrictions sémantiques, restrictions qui ne sont pas explicitées non plus. Comme nous allons le voir avec les travaux de (Asher *et al.*, n.d.), les variations de quantification, lorsqu'elles sont inter-

prêtées hors contexte, peuvent être expliquées sur la base des informations encodées dans les items en présence, et cela sans faire appel au contexte, toujours si difficile à définir.

3.2.3. Asher et Pustejovsky (AP) : une vision relationnelle du type pointé

(AP) visent aussi une résolution formelle des conflits de types, mais ils le font en agissant directement sur les individus, c'est-à-dire sur le nombre et le type des variables élémentaires de la formule courante. Dans leur approche, un type pointé est non seulement vu comme un type complexe liant deux types simples, mais est de plus défini par deux relations entre l'individu (ou l'ensemble d'individus) de type pointé et chacun des individus de type simple (ou les ensembles d'individus correspondants). Plus formellement, un type pointé, noté $\sigma_1 \bullet \sigma_2$ est bien formé si σ_1 et σ_2 sont eux-mêmes des types bien formés. De plus, un type pointé de la forme $\sigma_1 \bullet \sigma_2$ n'est valide que si les relations suivantes sont vraies de x , de type $\sigma_1 \bullet \sigma_2$, y , de type σ_1 et z , de type σ_2 :

$$\begin{aligned} (O_Elaboration_{\sigma_1}(x, y) \wedge x : \sigma_1 \bullet \sigma_2) &\rightarrow y : \sigma_1 \\ (O_Elaboration_{\sigma_2}(x, z) \wedge x : \sigma_1 \bullet \sigma_2) &\rightarrow z : \sigma_2 \end{aligned}$$

Ces relations peuvent être vues dans le lexique comme le reflet de nos connaissances encyclopédiques à propos de la classe d'objets désignée par un item à type pointé et du point des calculs effectués, comme le moyen d'extraire l'une des facettes sémantiques dans le sens complexe d'un polysème logique. Un nom comme *livre* est considéré par nombre d'auteurs comme désignant une classe d'objets complexes à deux facettes sémantiques, l'une physique, l'autre informationnelle. Les relations dans (AP) visent à représenter formellement le lien existant entre la classe d'objets complexes et chacune de ses facettes sémantiques. Pour cette raison, nous assimilons la définition du type pointé à une approche *relationnelle*.

Le nom *livre* est de type $p \bullet i \rightarrow t$ et il est représenté par le terme $\lambda x.livre(x)$. De plus, le type pointé $p \bullet i$ n'est valide que si les deux relations *OEP* (abréviation de *Object_Elaboration_Phys*) et *OEI* (abréviation de *Object_Elaboration_Info*) sont vérifiées et sont telles que :

$$\begin{aligned} (Object_Elaboration_Phys(x, y) \wedge x : p \bullet i) &\rightarrow y : p \\ (Object_Elaboration_Info(x, z) \wedge x : p \bullet i) &\rightarrow z : i \end{aligned}$$

La résolution des conflits de types dus à l'occurrence d'une variable de type pointé passe par des règles d'inférences particulières. Celles-ci (au nombre de 6) formalisent deux méthodes d'accommodation. Deux paires de règles *exploitent* les types pointés et font apparaître *localement* le type simple manquant dans le conflit à résoudre. Ces règles sont en fait dédoublées selon qu'elles exploitent un type pointé $\sigma_1 \bullet \sigma_2$ sur le type simple de gauche, σ_1 , ou sur le type simple de droite, σ_2 . Les deux dernières règles *introduisent* un type pointé, soit à partir du type simple de gauche, soit à partir de celui de droite.

Bien que différentes, ces règles agissent toutes directement au niveau des individus, c'est-à-dire au niveau des variables élémentaires dans la formule courante. L'introduction d'un nouveau type, que ce soit par exploitation ou par introduction, impose l'introduction d'une nouvelle variable en même temps que l'introduction de la relation *Object_Elaboration* appropriée étant donné les contraintes de types en présence. Par conséquent, la taille des formules augmente au fur et à mesure des résolutions de conflits. Enfin, dans le système défini par (AP), les règles d'inférence se distinguent des règles d'inférence classiques sur deux points. Premièrement, les expressions de types appartenant normalement à l'ensemble des hypothèses sont représentées par des contextes pouvant être transformés par élimination, ajout ou substitution d'une expression de type. Deuxièmement, les conflits de types sont traités au niveau des variables élémentaires et non plus au niveau des variables fonctionnelles.

Dans ce cadre, le conflit entre *livre* et *rouge* dans la phrase *Marie a traduit le livre rouge* ne sera pas directement pris en compte au niveau des prédicats, c'est-à-dire dans l'application entre le terme correspondant à l'épithète *rouge*, $\lambda Q \lambda v.rouge(v) \wedge Q(v)$, de type $(p \rightarrow t) \rightarrow p \rightarrow t$, et le terme correspondant au nom *livre*, $\lambda u.livre(u)$, de type $p \bullet i \rightarrow t$. Au contraire, le conflit est pris en compte au niveau des individus car celui qu'on voit apparaître au niveau des prédicat n'est pas considéré comme bloquant. Ainsi, le conflit est pris en compte dans l'application $\lambda v.rouge(v) \wedge (\lambda u.livre(u))(v)$ avec le contexte $\langle x : p \bullet i, v : p \rangle$. Sa résolution passe par l'utilisation de la règle d'exploitation à gauche ci-dessous¹⁵. Le but de cette règle est de retyper la variable t en la faisant passer du type $\alpha \bullet \beta$ au type α attendu par la variable x . Pour cela, la règle introduit une variable existentiellement quantifiée y , de type $\alpha \bullet \beta$ qui va se substituer à toute occurrence de t dans ϕ .

$$\frac{\lambda t.\phi[x], c(t : \alpha \bullet \beta, x : \alpha) \quad (S(a, b) \wedge a : \alpha \bullet \beta) \rightarrow b : \alpha}{\lambda t.\exists y(S(y, t) \wedge \phi[x]), c'(y : \alpha, x : \alpha \bullet \beta, t : \alpha)}$$

Les conflits entre prédicats n'étant pas considérés comme bloquants dans ces travaux, l'application entre le terme correspondant à *rouge* et celui correspondant à *livre* aboutit au terme $\lambda v.rouge(v) \wedge (\lambda u.livre(u))(v)$ avec le contexte $\langle u : p \bullet i, v : p \rangle$. Pour utiliser la règle d'exploitation gauche ci-dessus, la variable λ -abstraite t est associée à u , ϕ est associée au terme $\lambda u.livre(u)$ et x est associée à la variable

15. Les règles d'exploitation sont considérées par leurs auteurs comme une formalisation plus aboutie du principe de coercion de type tel qu'il est représenté dans (Pustejovsky, 1995). Pour une vision plus synthétique, nous avons simplifié la règle initiale et y avons intégré une seconde étape du calcul qui permet d'assigner à la nouvelle variable y le type dicté selon la règle *Object_Elaboration* appropriée.

Notations particulières : (1) la forme $\lambda t.\phi[x]$ représente une application entre une formule ϕ au sein de laquelle au moins une occurrence de la variable t , de type $\alpha \bullet \beta$, est λ -abstraite et le sous-terme x qui est une variable de type α . Le conflit résolu par cette règle se situe donc entre les variables élémentaires t et x ; c et c' représentent des contextes d'assignations de types et désignent respectivement le contexte initial et le contexte final résultant de l'application de la règle; la forme $(S(a, b) \wedge a : \alpha \bullet \beta) \rightarrow b : \alpha$ représente la relation d'élaboration liant le type simple de gauche, α , à son type pointé $\alpha \bullet \beta$.

v . Si, comme le propose cette règle, on substitue dans ϕ toute occurrence de u par la nouvelle variable y , introduite par la règle, existentiellement quantifiée et de type $p \bullet i$, alors ϕ devient $\lambda u. \exists y \text{ OEP}(y, u) \wedge \text{livre}(y)$ avec le contexte modifié $\langle y : p \bullet i, u : p \rangle$. Le conflit étant résolu entre la variable λ -abstraite de ϕ et la variable v apportée par le cotexte (ici l'adjectif *rouge*), une β -réduction conduit au terme $\lambda v. \exists y \text{ rouge}(v) \wedge \text{OEP}(y, v) \wedge \text{livre}(y)$ avec le contexte $\langle y : p \bullet i, v : p \rangle$.

Le second conflit rencontré se situe entre le syntagme *le livre rouge* et le verbe *traduire*. Dans le terme correspondant, $\exists v \exists y (\text{rouge}(v) \wedge \text{OEP}(y, v) \wedge \text{livre}(y) \wedge (\lambda u \lambda w. \text{traduit}(w, u))(v))$ avec le contexte $\langle v : p, y : p \bullet i, u : i, w : \text{hum} \rangle$, il apparaît un conflit entre la variable u de type i (c'est le patient de *traduire*) et la variable v de type p ¹⁶. La résolution de ce conflit suit une procédure particulière qui consiste en une double introduction du type pointé.

La règle d'introduction à partir du type de gauche est reproduite ci-dessous dans une version simplifiée¹⁷.

$$\frac{\Psi[\lambda t. \phi[x]], c(x : \alpha, t : \alpha \bullet \beta) \quad (S(a, b) \wedge a : \alpha \bullet \beta) \rightarrow b : \alpha}{\Psi_x^{\exists v} [S(x, y)][\lambda t. \phi[x]], c'(x : \alpha \bullet \beta, y : \alpha, t : \alpha \bullet \beta)}$$

Comme le montre la notation $\Psi[\lambda t. \phi[x]]$, cette règle s'applique à l'ensemble de la formule courante Ψ dans laquelle le type α est associé à la variable x . La notation $\Psi_x^{\exists v}$ indique que partout, sauf dans le terme $\lambda t. \phi[x]$, toutes les occurrences de x sont remplacées par une occurrence de la nouvelle variable existentiellement quantifiée y qui reçoit l'ancien type de x , le type α . Enfin, la variable x garde sa quantification initiale dans Ψ mais est retypée par le type pointé.

Dans le conflit entre *traduire* et *le livre rouge*, pour utiliser cette règle, on établit les égalités $[\Psi = \exists v \exists y (\text{rouge}(v) \wedge \text{OEP}(y, v) \wedge \text{livre}(y) \wedge (\lambda u \lambda w. \text{traduit}(w, u))(v))]$ et $[\lambda t. \phi[x] = \lambda u \lambda w. \text{traduit}(w, u)(v)]$. L'application de cette règle conduit à l'introduction d'une nouvelle variable z , existentiellement quantifiée et du type de v , le type p . Cette nouvelle variable z remplace ensuite toutes les occurrences de v dans toute la formule correspondant à Ψ sauf dans le sous-terme conflictuel correspondant à la représentation de *traduire*. Enfin, la variable v garde sa quantification existentielle, mais est retypée et donc associée au type $p \bullet i$, et la relation *OEP* est introduite et prédique les variables v et z . Le résultat de ces différentes opérations est montré dans le terme suivant :

$$\exists z \exists v \exists y (\text{rouge}(z) \wedge \text{OEP}(y, z) \wedge \text{livre}(y) \wedge \text{OEP}(v, z) \wedge (\lambda u \lambda w. \text{traduit}(w, u))(v))$$

avec le contexte modifié $\langle y : p \bullet i, v : p \bullet i, z : p, u : i, w : \text{hum} \rangle$.

Comme le montre le terme ci-dessus, le conflit persiste, mais implique cette fois la variable de type pointé v et la variable de type simple u . Selon (Asher *et al.*, n.d.), une

16. Nous précisons ici que la forme choisie pour représenter la composition du verbe *traduire* avec son complément *le livre rouge* ne tient pas compte du blocage de la variable w représentant l'agent de *traduire*.

17. Cette procédure n'étant pas explicitée précisément dans (Asher *et al.*, n.d.), ce qui suit n'est peut-être pas tout à fait cohérent avec l'intuition des auteurs.

seconde règle d'introduction, à droite cette fois, est appliquée pour faire apparaître le type pointé sur la variable conflictuelle et λ -abstraite u dans le terme représentant *traduire*. Le terme résultant est alors de la forme

$$\exists z \exists v \exists y (rouge(z) \wedge OEP(y, z) \wedge livre(y) \wedge OEP(v, z) \wedge (\lambda u \lambda w. \exists t OEI(u, t) \wedge traduit(w, t))(v))$$

avec le contexte modifié $\langle v : p \bullet i, z : p, y : p \bullet i, u : p \bullet i, w : hum, t : i \rangle$.

L'ensemble des conflits étant résolu, la formule correspondant au groupe verbal est de la forme

$$\lambda w. \exists t \exists z \exists v \exists y (rouge(z) \wedge OEP(y, z) \wedge livre(y) \wedge OEP(v, z) \wedge OEI(v, t) \wedge traduit(w, t))$$

avec le contexte $\langle v : p \bullet i, z : p, y : p \bullet i, w : hum, t : i \rangle$.

La formule ci-dessus représente un cas problématique pour les travaux de (AP) car elle demande une accommodation entre les types simples du type pointé de *livre*. La formule correspondant au groupe verbal *a traduit le livre rouge* comporte deux variables de type pointé et deux instances de la relation *OEP* partageant la même variable de type p . La question qui se pose est celle de l'identité ou de la différence entre les deux variables de type pointé. La seconde remarque concerne le nombre d'occurrences de la relation *OEP*. Ce nombre n'est pas cohérent avec le nombre de fois où le nom *livre* s'est trouvé en conflit avec une contrainte de type p . Enfin, la dernière remarque concerne les différentes possibilités d'application des règles. La résolution des conflits dans la phrase *Marie a traduit le livre rouge* peut aussi être obtenue avec deux autres façons d'appliquer les règles. La première conduit à un résultat identique et consiste simplement à remplacer la dernière introduction par une exploitation à droite. La seconde méthode conduit à une formule plus cohérente (une seule variable de type pointé et une seule occurrence des relations *Object_Elaboration* par type de conflit). Elle consiste à appliquer une règle d'introduction pour résoudre le conflit entre *livre* et *rouge*, suivie d'une règle d'exploitation à droite pour résoudre le second conflit. La formule résultant de cette seconde méthode est de la forme $\lambda w. \exists v \exists x \exists y (rouge(v) \wedge OEP(y, v) \wedge livre(y) \wedge OEI(y, x) \wedge traduit(w, x))$ avec le contexte $\langle w : hum, v : p, y : p \bullet i, x : i \rangle$.

3.3. Une représentation unifiée des travaux antérieurs

Les travaux de (Pinkal *et al.*, 2000) et (Asher *et al.*, n.d.) ont deux avantages qui sont préservés dans le traitement des polysèmes logiques de (Jacquey, 2004).

Le système de (PK) propose une méthode expressive et robuste qui permet d'accommoder les contraintes de types réunies dans une coprédication. Leur méthode a, en outre, l'avantage de garder la trace des différentes restrictions sélectionnelles qui sont en jeu. Cependant, la vision fonctionnelle du type pointé qu'ils développent demande une contextualisation des quantificateurs avec des coprédications quantifiées.

Le calcul de (AP) passe par ce que nous avons appelé une définition *relationnelle* des types pointés. Leur définition est relationnelle dans le sens où les variables de type pointé sont liées à leurs contreparties de type simple par les deux relations *Object_Elaboration*. Comme le lien entre les variables de type pointé et les variables de type simple est de nature relationnelle, et non pas de nature forcément fonctionnelle, le nombre d'individus de chacun des types en présence n'est pas forcément identique. Ainsi, *traduire les trois derniers livres de Chomsky* dénote en première interprétation trois objets informationnels, mais un nombre indéterminé d'objets de type pointé $p \bullet i$ et d'objets physiques.

Dans des cas moins complexes que celui de l'accommodation entre les types simples d'un type pointé, le calcul de (AP) conduit bien à des formules correspondant à des interprétations attendues sur le plan du sens qui supporte la quantification. Ainsi, la phrase *Marie a volé tous les livres de la bibliothèque* est représentée par une formule bien typée dans laquelle le sens physique de *livre* est quantifié universellement. Inversement, la formule représentant la phrase *Marie a lu tous les livres de la bibliothèque* fait bien porter la quantification universelle sur le sens complexe, symbolisé par le type pointé $p \bullet i$, mais dans ce cas, les gloses possibles sont fortement culturelles dans la mesure où elles s'appuient sur notre point de vue de ce qu'est l'acte de lecture. Une glose possible de cette interprétation pourrait donc être *tout objet livre dont il existe un support physique situé dans la bibliothèque a été tenu par Marie et a une contrepartie informationnelle dont Marie a ainsi pris connaissance*. Par conséquent, les difficultés rencontrées avec l'approche de (AP) concerne les aspects opératoires. Nous avons rappelé qu'il est possible de résoudre les conflits en appliquant les règles d'inférences de manière différente, cela conduisant d'une part à plus de formules bien formées que nécessaire, et d'autre part à des formules représentant des interprétations non attendues étant donné les éléments linguistiques en présence. Par exemple, la phrase *Marie a lu tous les livres de la bibliothèque* peut aussi être représentée par une formule, bien typée elle aussi, mais dans laquelle le sens physique de *livre* supporte la quantification universelle. Cette seconde interprétation revient à supposer que Marie a lu tous les livres de la bibliothèque, que ceux-ci comptent ou non plusieurs exemplaires d'une même œuvre.

À partir de ce constat, (Jacquey, 2004) propose une unification des travaux de (Asher *et al.*, n.d.) et (Pinkal *et al.*, 2000) sur le plan de la définition du type pointé et sur celui des règles d'accommodation. Nous la rappelons brièvement ici.

3.3.1. Définition des types pointés de la forme $\alpha \bullet \beta$

La modélisation de (Jacquey, 2004) unifie sans les confondre celles de (PK) et de (AP) : un type pointé est décrit à l'aide de fonctions partielles, fonctions correspondant à l'extraction des sous-termes dans (Pinkal *et al.*, 2000), ou bien ce même type est défini par le biais des deux relations conceptuelles reprises de (Asher *et al.*, n.d.). La distinction entre les deux approches à l'intérieur même de la définition des types pointés est préservée car elles relèvent de deux procédures différentes de résolution des conflits de types apparaissant avec les types pointés. Lorsqu'un conflit ne porte que

sur des paramètres λ -abstrait, il peut être vu comme un conflit entre des contraintes de types, ce conflit ne resurgissant pas sur le comptage des individus. Dans ce cas, les contraintes peuvent être unifiées comme le proposent (PK) selon une *accommodation fonctionnelle* des conflits de types. En revanche, lorsque le conflit porte sur au moins une variable quantifiée, sa résolution resurgit sur le comptage des individus. Dans ce cas, l'approche de (AP), dite *accommodation relationnelle* semble plus appropriée.

La *définition fonctionnelle* du type pointé, traduit le fait qu'un type simple est toujours accessible à partir d'un type pointé sous la forme de *fonctions de projection*, et pour cela s'appuie sur des fonctions de projection de la forme $\Pi_ \alpha$ et de type $\alpha \bullet \beta \rightarrow \alpha$ ¹⁸.

La *définition relationnelle* du type pointé traduit le fait qu'il existe une relation entre une entité dénotée par un polysème logique et une ou plusieurs entités dénotée(s) par l'un des sens simples. Ce fait est représenté par le biais de *prédicats relationnels*¹⁹, de forme $R\alpha$ et de type $\alpha \bullet \beta \rightarrow \alpha \rightarrow t$ ²⁰.

3.3.2. Deux sortes de résolution des conflits de types

L'*accommodation fonctionnelle*, utilisant la *définition fonctionnelle* du type pointé, vise, quel que soit le cas de figure rencontré, à conserver l'accessibilité du type pointé. Étant donné un type pointé $\alpha \bullet \beta$ appartenant à l'ensemble des types et γ un type quelconque de cet ensemble²¹, étant donné l'ensemble des termes contenant les fonctions de projections $\Pi_ \alpha$ et $\Pi_ \beta$, toute application (MN) sera de type $\alpha \bullet \beta \rightarrow \gamma$ quel que soit le type impliqué dans la construction du type fonctionnel du terme $M : \alpha, \beta$ ou $\alpha \bullet \beta$. Cette accommodation fonctionnelle repose conformément au cadre formel sous-jacent sur des schémas d'assignation de types.

Le schéma [2] permet de résoudre le conflit entre un terme M de type $(\beta \rightarrow t) \rightarrow \beta \rightarrow t$ et un terme N de type $\alpha \bullet \beta$. Un tel schéma serait utilisé dans la résolution d'un conflit entre *rouge*, de type $(p \rightarrow t) \rightarrow (p \rightarrow t)$, et *livre*, de type $p \bullet i \rightarrow t$.

$$\frac{A \vdash M : (\beta \rightarrow \gamma) \rightarrow (\beta \rightarrow \gamma) \quad A \vdash N : \alpha \bullet \beta \rightarrow \gamma}{A \vdash (MN) : \alpha \bullet \beta \rightarrow \gamma} \quad [2]$$

Les β -réductions correspondant à des schémas d'assignation de types tels [2] que modifient le terme M en y remplaçant les variables élémentaires conflictuelles par des termes complexes de la forme $\Pi_ type_ simple(x)$ correspondant à l'application

18. Ces fonctions sont une représentation dans (Jacquey, 2004) du rôle joué par la règle d'inférence dans (Pinalk *et al.*, 2000) qui contraint l'extraction des facettes.

19. Ces prédicats relationnels sont la traduction dans (Jacquey, 2004) des relations d'élaboration, *OEP* ou *OEI*, qui lient les variables de type simple dans (Asher *et al.*, n.d.).

20. Étant donné un type pointé de forme $\alpha \bullet \beta$, les éléments formels concernant sa partie droite β sont équivalents à ceux de α .

21. Le type γ est ici volontairement sous-spécifié. Dans tous les exemples commentés ici, ce type γ sera égal au type t , mais dans un souci de généralité, ce type est resté abstrait afin qu'il puisse être égal à un type fonctionnel, par exemple.

entre la fonction de projection adéquate et la variable conflictuelle. Dans le cas de *livre rouge*, le type simple entrant en conflit étant le type p , alors ce sera la règle Π_p qui sera utilisée. La règle [3] ci-dessous combine le schéma [2] avec la β -réduction.

$$\frac{(\lambda\phi^{\beta \rightarrow \gamma} \lambda x^\beta. (\Psi[\phi(x)])^\gamma)^{(\beta \rightarrow \gamma) \rightarrow (\beta \rightarrow \gamma)} \quad N^{\alpha \bullet \beta \rightarrow \gamma}}{\lambda x^{\alpha \bullet \beta}. (\Psi[(\Pi_\beta(x))][(\phi/N)(x)])^\gamma} \quad [3]$$

Cette règle indique comment un terme (MN) peut être correctement construit et typé dans le cadre d'une accommodation projective. Elle montre en particulier que le remplacement des variables élémentaires conflictuelles s'applique dans tout le terme M , équivalent à $\lambda\phi\lambda x. (\Psi[\phi(x)])$, sauf dans le sous-terme de celui-ci représenté par ϕ , c'est-à-dire la variable fonctionnelle conflictuelle²². Plus précisément, cette règle décrit deux transformations :

1) puisque c'est la variable élémentaire x qui est responsable du conflit entre la variable fonctionnelle ϕ , de type $\beta \rightarrow t$, et le terme rencontré N , de type $\alpha \bullet \beta \rightarrow t$, alors cette variable x est retypée par le type $\alpha \bullet \beta$ dans toute la formule Ψ , conduisant ainsi au retypage de la variable conflictuelle ϕ et neutralisant le conflit entre ϕ et N ;

2) pour neutraliser le conflit survenant dans Ψ à la suite de cette première transformation, la variable élémentaire x de type β est remplacée partout sauf dans ϕ par l'application entre la fonction de projection Π_β et la variable x retypée en $\alpha \bullet \beta$.

Dans le cas du syntagme *livre rouge*, représenté par l'application entre le terme $\lambda R \lambda x. R(x) \wedge rouge(x)$, de type $(p \rightarrow t) \rightarrow (p \rightarrow t)$, et le terme $\lambda y. livre(y)$, de type $p \bullet i \rightarrow t$, la règle correspondant à la règle [3], c'est-à-dire sur la partie gauche α du type pointé, a pour effet de remplacer toutes les occurrences conflictuelles de x , car de type p plus général que $p \bullet i$, dans le terme correspondant à *rouge* sauf dans le sous-terme R auquel on substitue le terme *livre*. À l'issue de cette première accommodation, l'application *livre rouge* est donc associée au terme $\lambda x. livre(x) \wedge rouge(\Pi_p(x))$ de type $p \bullet i \rightarrow t$.

Puisque l'accommodation fonctionnelle conserve le type pointé et puisque celle-ci ne porte jamais sur des variables quantifiées, l'accommodation relationnelle, qui implique au moins une variable quantifiée, sera toujours d'un type de la forme $(\alpha \bullet \beta \rightarrow \gamma) \rightarrow \gamma$ ²³. Quel que soit le cas de figure rencontré, le terme résultant d'une application où le conflit est résolu par une accommodation relationnelle doit être de type γ . Le schéma [4] s'applique lorsqu'un terme quantifié de type $(\alpha \bullet \beta \rightarrow \gamma) \rightarrow \gamma$

22. La méthode d'une application partielle d'une règle a été introduite dans (Asher *et al.*, n.d.). De plus, par rapport à ces travaux, il faut préciser que le résultat de la β -réduction ne peut être représenté par un sous-terme de la forme $[x/\Pi_{beta}(x)]$ car x n'a pas le même type dans ses deux occurrences : il est de type β dans la première et de type $\alpha \bullet \beta$ dans la seconde.

23. Le résultat d'une accommodation fonctionnelle étant toujours de type $\alpha \bullet \beta \rightarrow t$ et le terme représentant un quantificateur étant toujours de type $(\gamma \rightarrow t) \rightarrow (\gamma \rightarrow t) \rightarrow t$, la composition de ces éléments est toujours de type $(\alpha \bullet \beta \rightarrow t) \rightarrow t$ avec $\alpha \bullet \beta$ un sous-type de γ .

rencontre un terme N dénotant des entités du type simple α , donc de type fonctionnel $\alpha \rightarrow \gamma$.

$$\frac{A \vdash M : (\alpha \bullet \beta \rightarrow \gamma) \rightarrow \gamma \quad A \vdash N : \alpha \rightarrow \gamma}{A \vdash (MN) : \gamma} \quad [4]$$

La β -réduction correspondant à ce schéma est plus complexe que dans le cas des accommodations projectives. Elle utilise un outil introduit par (AP), symbolisé par $\Psi_{x^{\exists y}}(R_type_simple(y, x))(\phi(x))$ et résumant trois mécanismes illustrés dans la règle [5] :

- 1) une nouvelle variable y , existentiellement quantifiée et de type pointé, est introduite et remplace toutes les instances de x dans la formule M sauf dans la sous-formule $\phi(x)$;
- 2) la relation $R_\alpha(y, x)$ est introduite dans la formule pour lier cette nouvelle variable y à la variable x considérée comme conflictuelle ici ;
- 3) à la suite du retypage de la variable x , la variable fonctionnelle conflictuelle ϕ est elle aussi retypée en fonction du terme N rencontré.

$$\frac{(\lambda \phi^{\alpha \bullet \beta \rightarrow \gamma} . (\Psi[\phi(x)])^\gamma)^{(\alpha \bullet \beta \rightarrow \gamma) \rightarrow \gamma} \quad N^{\alpha \rightarrow \gamma}}{\exists y^{\alpha \bullet \beta} . (\Psi[R_\alpha(y, x^\alpha)][(\phi^{\alpha \rightarrow \gamma} / N)(x^\alpha)]^\gamma)} \quad [5]$$

3.4. Bilan

Cette modélisation de (Jacquey, 2004) satisfait les contraintes fixées au départ : les propriétés des polysèmes logiques sont traitées, le traitement se situe au niveau strictement lexical, la disjonction est évitée, chaque sens des polysèmes, représentés chacun par un type, est accessible indépendamment des autres et les interprétations calculées dans ce cadre formel sont conformes aux interprétations observées.

Nous abordons maintenant une évaluation qualitative de cette modélisation en montrant qu'elle peut être adaptée pour rendre compte du contenu sémantique des noms d'action ambigus entre processus et artefact. Rappelons que ces noms ne partagent pas totalement les propriétés des polysèmes logiques prototypiques : la copré-dication est plus contrainte (elle est difficilement acceptable lorsqu'elle est basée sur la coordination) et l'ambiguïté processus/artefact n'apparaît pas dans toutes les configurations argumentales et aspectuelles possibles pour ces noms.

4. Extensions et spécifications de la modélisation de la polysémie logique pour les noms d'action ambigus entre processus et artefact

Les extensions ajoutées ont pour but de prendre en compte les réalisations syntaxiques particulières des syntagmes dans lesquels apparaissent les noms d'action, c'est-à-dire des groupes nominaux complexes. Ce point particulier est traité par le biais d'une modélisation des prépositions.

De plus, l'ensemble des types et celui des termes doivent être étendus afin d'intégrer les éléments particuliers impliqués dans la définition d'un nouveau type pointé *proc • art* pour cette sous-classe de polysèmes logiques. En effet, à partir de la définition d'un type pointé (section 3.2), l'existence du type pointé *proc • art* implique la création des deux fonctions de projection correspondantes Π_{proc} qui permet de construire un terme de type *proc* (sens processif) et Π_{art} qui fait de même pour un terme de type *art* (sens résultatif). Ensuite, afin de permettre l'accommodation relationnelle, les prédicats relationnels R_{proc} et R_{art} sont créés.

4.1. Propriétés syntaxiques des groupes nominaux complexes et représentation des prépositions

L'une des caractéristiques syntaxiques majeures des groupes nominaux complexes est l'optionalité de certains de leurs arguments. Cette propriété ne se vérifie pas uniquement sur les noms d'action ambigus tels que *construction* ou *reproduction*, mais aussi sur des noms d'états comme *connaissance* ou sur des noms à interprétations multiples comme *livre*.

Dans la mesure où la question du mécanisme d'assemblage des différents constituants d'un groupe nominal complexe n'est pas une caractéristique propre aux noms d'action ambigus, nous ferons simplement une hypothèse. Nous supposons dans les prochains calculs que le mécanisme d'assemblage des constituants prépositionnels internes à un groupe nominal passe par l'utilisation des prépositions comme une « glue » sélective entre les constituants. Les groupes nominaux complexes, dont la tête syntaxique est un nom d'action, peuvent être de plusieurs formes.

la construction de la maison par les ouvriers	$Det\ N\ de_NP_1\ par_NP_0$
la construction de la maison	$Det\ N\ de_NP_1$
la construction du meilleur élève	$Det\ N\ de_NP_0$
la construction	$Det\ N$

Toutes ces formes doivent être traduites par des λ -termes de type, ou d'un sous-type de, $e \rightarrow t$ pour pouvoir être composées avec leur déterminant. En appliquant successivement les règles de typage de base du λ -calcul, application et abstraction, il est possible de déduire le type de toute préposition : $((e \rightarrow t) \rightarrow t) \rightarrow (e \rightarrow t) \rightarrow e \rightarrow t$.

4.2. Propriétés sémantiques des noms d'action ambigus contraintes par leurs prépositions

Comme précisé dans la section (2.3), les différentes configurations dans la réalisation des arguments conditionnent les interprétations des noms d'action. Ces interprétations peuvent désigner un processus ou un artefact, ou être ambiguës. Ce sont les types associés à chacune des prépositions introduisant les arguments qui sont chargées de coder les différents types d'interprétations non ambiguës²⁴.

- La préposition *par*, étant caractéristique d'une interprétation processive lorsqu'elle introduit l'agent de l'action correspondante, du moins dans les exemples discutés ici, porte sur une variable individuelle de type *proc*. Elle est représentée par une fonction trivalente : l'agent, le patient et un individu de type *proc*. Cette fonction est un terme de la forme $[\lambda P \lambda Q \lambda x. (P)(\lambda u. agent(u, x)) \wedge Q(x)]$ et de type $((hum \rightarrow t) \rightarrow t) \rightarrow (proc \rightarrow t) \rightarrow proc \rightarrow t$. La représentation de l'agent $\lambda u. agent(u, x)$ est appliquée au terme qui sera substitué à la variable fonctionnelle P lors du calcul. L'agent désignant par défaut un être humain, la variable P doit être de type $(hum \rightarrow t) \rightarrow t$. Par ailleurs, la variable fonctionnelle Q sera remplacée par la représentation du nom d'action. Puisque l'interprétation de la préposition *par* n'est pas ambiguë dans ce cas, alors la variable Q est de type $proc \rightarrow t$.

- La préposition *de* regroupe deux emplois différents. Lorsqu'elle introduit l'objet d'un nom de création tel que *construction*, objet qui joue le rôle de patient, elle est caractéristique d'une interprétation processive et à ce titre, introduit un groupe nominal dont le nom recteur est de type $proc \rightarrow t$. Cette possibilité serait réalisée dans un groupe nominal comme *la construction d'un cocon_{patient}*, construit à partir de l'exemple (5c).

En revanche, elle est ambiguë lorsqu'elle introduit le patient d'un nom tel que *reproduction*. Cette possibilité serait réalisée dans un groupe nominal comme *une reproduction des séances publiques_{patient}*, construit à partir de l'exemple (5a). Elle introduit alors un groupe nominal dont le nom recteur est de type $proc \bullet art \rightarrow t$. De plus, cette seconde interprétation possible de la préposition *de* apparaît aussi, avec les deux types de noms *construction* et *reproduction*, lorsqu'elle introduit un agent en *de* : *les constructions des animaux* issu de l'exemple (5c) et *les reproductions des rédacteurs* issu de l'exemple (5a).

Par conséquent, deux prépositions *de* sont définies.

24. Cette solution, quoique certainement un peu trop brutale du fait du petit nombre de facteurs pouvant être pris en compte, s'appuie sur l'hypothèse qu'un groupe nominal régi par un nom d'action fonctionne comme une phrase à l'infinif. Dans le cadre phrastique, le rôle des constituants est partiellement reconnaissable en fonction de leur position syntaxique et en fonction de leurs sens lexicaux. De la même manière, cette approche des prépositions est fondée sur l'hypothèse que celles-ci, en coopération avec les sens lexicaux des groupes qu'elles lient entre eux, sont à même de rendre compte partiellement de leurs rôles respectifs. Cette hypothèse de similitude entre les deux fonctionnements explique aussi l'introduction des rôles thématiques *agent* et *patient*.

- La préposition de_1 introduit l'objet d'un nom de création comme *construction* dans un \overline{N} de la forme $N_de_NP_1$. Elle est caractéristique de l'interprétation processive du nom N . Elle correspond donc à un terme de la forme $[\lambda P \lambda Q \lambda x. (P)(\lambda u. patient(u, x)) \wedge Q(x)]$ et de type $((art \rightarrow t) \rightarrow t) \rightarrow (proc \rightarrow t) \rightarrow proc \rightarrow t$.

- Dans un \overline{N} de la forme $N_de_NP_1$, la préposition de_2 peut introduire, ou bien le patient ou l'agent d'un nom comme *reproduction*, ou bien l'agent d'un nom comme *construction*. Elle est caractéristique d'une interprétation ambiguë du nom d'action N qui porte donc sur deux aspects : (1) l'interprétation du nom d'action qui peut être aussi bien processive qu'artefactuelle et (2) le rôle joué par son objet, uniquement agent avec des noms comme *construction*, ou agent ou patient, avec des noms comme *reproduction*. L'ambiguïté sur l'interprétation du nom d'action est représentée à l'aide du type pointé $proc \bullet art$ et l'ambiguïté sur le rôle joué par l'objet de celui-ci est représentée par la sous-détermination du rôle via l'utilisation du terme *actant* dans le terme représentant la préposition. Elle est donc représentée par un terme de la forme $[\lambda P \lambda Q \lambda x. (P)(\lambda u. actant(u, \Pi_{proc}(x))) \wedge Q(x)]$ et de type $((\alpha \rightarrow t) \rightarrow t) \rightarrow (proc \bullet art \rightarrow t) \rightarrow proc \bullet art \rightarrow t$. Ce dernier emploi de *de* appelle un commentaire. La fonction $\lambda u. actant(u, \Pi_{proc}(x))$ est de cette forme car elle attend un argument de n'importe quel type α . Cette fonction signifie que, étant donné u le patient (et/ou l'agent) et x le référent d'un nom d'action ambigu, u est tel qu'il vérifie le prédicat *actant* lorsque ce prédicat est appliqué à l'interprétation processive du nom d'action, interprétation qui est notée par le terme $\Pi_{proc}(x)$. Comme x est de type $proc \bullet art$, cette variable ne peut dénoter, à elle seule, un événement. C'est la fonction de projection Π_{proc} qui, une fois appliquée à x , permet de désigner un référent événementiel.

4.3. Application à quelques cas d'étude

Les éléments nécessaires au traitement des exemples *Une reproduction des séances publiques qui a été envoyée directement à l'imprimeur a été effectuée selon les règles prescrites par l'Assemblée* et *Toutes les reproductions des séances publiques directement envoyées à l'imprimeur sont effectuées selon les règles prescrites par l'Assemblée* ne demandent pas d'explications particulières. Nous donnons leur représentation et leur type dans les exemples suivants²⁵.

- (9) a. *un/une*
 b. $\lambda P \lambda Q. \exists x (P(x) \wedge Q(x))$
 c. $(\alpha \rightarrow t) \rightarrow (\alpha \rightarrow t) \rightarrow t$

25. La relative *qui a été envoyée directement à l'imprimeur* est simplifiée sous une forme adjectivale, *envoyée*, et associée au prédicat *envoyé_à_imprimeur*. De même, le groupe verbal *sont effectuées selon les règles prescrites par l'Assemblée* est simplifié sous la forme du prédicat *effectué*.

- (10) a. *tous les*
 b. $\lambda P \lambda Q. \forall x (P(x) \rightarrow Q(x))$
 c. $(\alpha \rightarrow t) \rightarrow (\alpha \rightarrow t) \rightarrow t$
- (11) a. *envoyé à l'imprimeur*
 b. $\lambda R \lambda x. \text{envoyé_à_imprimeur}(x) \wedge R(x)$
 c. $(art \rightarrow t) \rightarrow art \rightarrow t$
- (12) a. *séances publiques*
 b. $\lambda x. \text{séance}(x)$
 c. $proc \rightarrow t$
- (13) a. *effectué selon les règles prescrites par l'Assemblée*
 b. $\lambda x. \text{effectué}(x)$
 c. $proc \rightarrow t$
- (14) a. *reproduction*
 b. $\lambda x. \text{reprod}(x)$
 c. $proc \bullet art \rightarrow t$

Construction du groupe *reproduction des séances publiques*

Ce groupe correspond à la structure parenthésée ((*de (le séance) reproduction*)). Dans la première étape, $M = de_2 = \lambda R \lambda S \lambda y. (R)(\lambda u. \text{patient}(u, \Pi_{proc}(y))) \wedge S(y)$ et $N = \text{les séances publiques} = \lambda Q. \exists x (\text{séance}(x) \wedge Q(x))$.

La construction du groupe *reproduction des séances publiques* est obtenue par une succession de règles d'application sans qu'aucun conflit n'apparaisse. Le terme obtenu, de type $proc \bullet art \rightarrow t$, est de la forme suivante :

$$(15) \lambda y. \exists x (\text{séance}(x) \wedge \text{patient}(x, \Pi_{proc}(y))) \wedge \text{reprod}(y)$$

Le type de ce terme, en particulier le type pointé de la variable λ -abstraite y encode bien le fait que le groupe nominal est ambigu.

Un exemple de coprédication

Dans la phrase *La reproduction des séances publiques qui a été envoyée directement à l'imprimeur a été effectuée selon les règles prescrites par l'Assemblée*, le groupe ambigu *reproduction des séances publiques* doit être accommodé selon deux contraintes de types : la relative, qui est de type $(art \rightarrow t) \rightarrow art \rightarrow t$ et le groupe verbal, qui est de type $proc \rightarrow t$.

La première contrainte à entrer en jeu est celle de la relative, traitée ici comme un adjectif épithète modifiant des artefacts. Le premier conflit n'impliquant aucune

variable quantifiée, la règle d’accommodation fonctionnelle [3] est utilisée.

Avec $[\phi(x) = R(z)]$ dans $[\Psi = \text{envoyé_à_imprimeur}(z) \wedge R(z)]$ et $[N = \lambda y. \exists x. (\text{séance}(x) \wedge \text{patient}(x, \Pi_{proc}(y)) \wedge \text{reprod}(y))]$, suite à la substitution dans Ψ de R par T , le terme obtenu est de type $(proc \bullet art \rightarrow t) \rightarrow t$ et de la forme :

$$(16) \lambda Q. \exists z \exists x (\text{envoyé_à_imprimeur}(\Pi_{art}(z)) \wedge \text{reprod}(z) \wedge \text{séance}(x) \wedge \text{patient}(x, \Pi_{proc}(z)) \wedge Q(z))$$

La seconde contrainte de types est celle apportée par le groupe verbal *a été effectuée selon les règles prescrites par l’Assemblée*. Celle-ci correspond au terme $\lambda u. \text{effectué}(u)$ qui est de type $proc \rightarrow t$. Un second conflit de type est donc détecté avec la variable quantifiée z , induisant l’utilisation de la règle d’accommodation relationnelle à gauche [5].

Avec $[\phi(x) = Q(z)]$ dans $[\Psi[\phi(x)] = \exists z \exists x (\text{envoyé_à_imprimeur}(\Pi_{art}(z)) \wedge \text{reprod}(z) \wedge \text{séance}(x) \wedge \text{patient}(x, \Pi_{proc}(z)) \wedge Q(z))]$, et avec $[N = \lambda u. \text{effectué}(u)]$, le terme obtenu, de type t , est de la forme :

$$(17) \exists z \exists x \exists y (\text{envoyé_à_imprimeur}(\Pi_{art}(y)) \wedge \text{reprod}(y) \wedge \text{séance}(x) \wedge \text{patient}(x, \Pi_{proc}(y)) \wedge R_{proc}(y, z) \wedge \text{effectué}(z))$$

Cette représentation est un terme correctement typé au sein duquel une nouvelle variable y existentiellement quantifiée joue le rôle de la variable initiale z . La nouvelle variable z est retypée de $proc \bullet art$ vers $proc$ et l’insertion du prédicat relationnel $R_{proc}(y, z)$ permet de lier entre elles la variable z retypée et la variable y nouvellement introduite.

De plus, les restrictions sélectionnelles de l’ensemble des prédicats logiques correspondent bien aux restrictions sélectionnelles des items lexicaux. Les conflits sont résolus sans perdre la trace de ces restrictions.

Coprédication et quantification

La phrase – *Toutes les reproductions des séances publiques directement envoyées à l’imprimeur sont effectuées selon les règles prescrites par l’Assemblée* – illustre ces deux propriétés des noms d’action ambigus. Puisqu’il s’agit de la même phrase que précédemment, excepté pour ce qui concerne le type de quantification, nous donnons simplement le terme obtenu à l’issue des mêmes étapes de calcul que précédemment. Ce terme, de type t , est de la forme :

$$(18) \forall z \exists x \exists y (\text{envoyé_à_imprimeur}(\Pi_{art}(y)) \wedge \text{reprod}(y) \wedge \text{séance}(x) \wedge \text{patient}(x, \Pi_{proc}(y)) \wedge S_{proc}(y, z) \rightarrow \text{effectué}(z))$$

5. Conclusion

Cet article a permis de montrer que la modélisation de (Jacquey, 2004) peut être adaptée aux noms d’action ambigus entre processus et artefact moyennant deux extensions : une modélisation des prépositions à la fois sur le plan de leur type et de leur

sens, et la définition d'un type pointé propre à cette sous-classe de polysémie logique. Cette adaptation montre que la modélisation initiale possède un degré d'expressivité qui va au-delà des cas prototypiques de polysémie logique. D'autre part, la perspective d'étendre la modélisation initiale des cas prototypiques de polysèmes logiques a été l'occasion d'aborder la question de la délimitation de ce phénomène sur le français. Comme nous l'avons mentionné plus haut, si l'on s'appuie sur l'hypothèse de (Kleiber, 1999) qui rapproche la polysémie logique de la métonymie, on s'aperçoit en interrogeant le TLF que 12,4 % de ses noms, soit 6 000 noms, peuvent être considérés comme relevant potentiellement de la polysémie logique. Dans cet ensemble, la classe de l'ambiguïté processus/artefact était intéressante à modéliser car il s'agit d'une des classes de polysémie logique la plus éloignée des cas prototypiques. En particulier, l'ambiguïté processus/artefact fait intervenir des types très différents, certainement plus que les types *physique* et *informationnel* avec le nom *livre* car avec cette dernière ambiguïté, la différence de types se situe entre des types d'objets, pas entre le domaine des objets et le domaine des événements. D'autre part, l'ambiguïté processus/artefact n'est pas aussi systématique que l'ambiguïté physique/informationnel de *livre*. Son apparition est contrainte par les réalisations argumentales des noms d'action concernés. Malgré ces différences importantes, un mécanisme sémantique commun, représenté par la modélisation proposée ici, semble pouvoir être appliqué aux deux classes.

Cependant, plusieurs questions restent en suspens. En particulier, les extensions proposées ne permettent pas en l'état d'exclure les coprédications bâties sur la coordination. En l'état actuel, en effet, les règles d'accommodation de types n'intégrant aucune contrainte d'ordre syntaxique, elles ne permettent pas de faire la différence entre une accommodation mise en œuvre dans une coprédication qui repose sur une relation anaphorique et une accommodation requise par une coprédication reposant sur une coordination.

Relativement à la question de la couverture sémantique offerte par cette modélisation, une évaluation qualitative des noms d'action facilement repérables dans le TLF (CNRS, 1976-1994) montre que beaucoup d'entre eux possèdent au moins un troisième sens, celui de l'activité en général, *la construction de chalets*, *la construction en climat difficile*, etc. Ce troisième sens, très courant avec les noms d'action, n'est actuellement pas représenté. Son intégration dans l'ensemble des types pose deux questions. Premièrement, le type *activ* qui serait utilisé pour symboliser ce sens d'activité doit-il être ajouté dans le type pointé *proc*•*art* au même niveau que *proc* et *art* ou le type *proc* est-il lui-même un type complexe contenant un type *accomplissement* et un type *activ*, et si oui, ce type complexe est-il un type pointé ou bien un type d'une autre nature ? Deuxièmement, dans tous les cas, se pose la question de la validité d'un type pointé composé, soit de plus de deux types simples, soit d'un type pointé et d'un type simple.

6. Bibliographie

- Asher N., Pustejovsky J., « The Metaphysics of Words in Context », n.d.
<http://www.dla.utexas.edu/depts/philosophy/faculty/asher/papers/MWC.ps>.
- CNRS, *TLF, Dictionnaire de la langue du 19ème et 20ème siècle*, CNRS-Gallimard, Paris, 1976-1994. 16 vol.
- Copestake A., Briscoe T., « Semi-productive Polysemy and Sense Extension », *Journal of Semantics*, vol. 12, p. 15-67, 1995.
- Cruse D., *Lexical Semantics*, Cambridge : Cambridge University Press, 1986.
- Cruse D., « Polysemy and related phenomena from a cognitive linguistic viewpoint », *Saint-Dizier, P. and Viegas, E. (eds.)*, 1995.
- Dowty D., « Tenses, times-adverbs and compositional semantic theory », *Linguistics and Philosophy*, vol. 5, p. 23-55, 1982.
- Dowty D., Wall R., Peters S., *Introduction to Montague Semantics*, Dordrecht, 1985.
- Godard D., Jayez J., « Types Nominaux et Anaphores ; le cas des objets et des événements », *Anaphores temporelles et (in-)cohérence*, Cahiers Chronos, 1996.
- Grimshaw J., *Argument Structure*, MIT Press, 1990.
- Hindley J., Seldin J., *Introduction to Combinators and λ -calculus*, vol. 1, London Mathematical Society, Student texts, 1986.
- Jacquey E., Ambiguïtés lexicales et traitement automatique des langues : Modélisation de la polysémie logique et application aux déverbaux d'action ambigus en français, PhD thesis, LORIA, Université de Nancy2, UFR Mathématiques et Informatique, 2001.
- Jacquey E., « Ambiguïté lexicale et quantification : une modélisation de la polysémie logique », in C. Corblin, Francis et Gardent (ed.), *Interpréter en contexte*, Hermes, 2004.
- Kleiber G., *Problèmes de sémantique : la polysémie en questions*, sens et structures, Villeneuve d'Ascq : Presses Universitaires du Septentrion, 1999.
- Milner J., *Ordres et raisons de langue*, Paris, Seuil, 1982.
- Pinkal M., Kohlhase M., « Feature Logic for Dotted Types : A Formalism for Complex Words Meanings », *ACL*, 2000.
- Pustejovsky J., « The Generative Lexicon », *Computational Semantics*, vol. 17, n° 4, p. 409-441, 1991.
- Pustejovsky J., *The Generative Lexicon*, MIT Press, 1995.
- Pustejovsky J., « The Semantics of Complex Types », *Langue Française*, 1996.
- Pustejovsky J., Anick P., « On The Semantic Interpretation of Nominals », *Proceedings of COLING*, Budapest, 1988.
- Samvelian P., Les nominalisations en français : arguments syntaxiques et actants sémantiques, PhD thesis, Université de Paris 7 - Denis Diderot, UFR de linguistique, TALANA, 1995.
- Vendler Z., *Linguistics in Philosophy*, Cornell University Press, chapter Verbs and Times, 1967.
- Verkuyl H., « Aspectual classes and aspectual composition », *Linguistics and Philosophy*, vol. 12, p. 39-94, 1989.