

Peut-on *couper* à la polysémie verbale ?

Françoise Gayral, Patrick Saint-Dizier

LIPN
Université Paris-Nord
F-93430 Villetaneuse
fg@lipn.univ-paris13.fr,

IRIT-CNRS
118, route de Narbonne
F-31062 Toulouse Cedex
stdizier@irit.fr

Résumé

Dans ce bref document, nous présentons des résultats préliminaires d'une méthode de description de la sémantique des formes prédicatives dans un cadre génératif. Nous proposons une méthode pour découper les sens, en identifiant les types d'inférences qu'ils entraînent. Nous proposons une analyse intégrée des métaphores et des métonymies, ainsi qu'une représentation des sens sous forme typée et sous-spécifiée.

Mots clés: sémantique lexicale, prédicats, lexique génératif.

1. Introduction

Si de nombreux travaux récents en sémantique lexicale et de la proposition traitent des variations de sens des noms avec l'étude de phénomènes comme la métonymie, la coertion (Pustejovsky 91, Lakoff et al. 80, Nunberg et al. 92, Ostler et al. 92, Pernelle 98, etc.), moins fréquente est l'étude de la polysémie verbale (Pustejovsky 95). Pourtant, le verbe est souvent considéré comme jouant un rôle particulier dans le processus de composition sémantique: élément organisateur, 'pivot' de la proposition, imposant certaines contraintes au reste de l'énoncé.

Cet article aborde quelques contours d'une méthode pour l'analyse et la modélisation de la variation de sens des termes prédicatifs en contexte. Les aspects liés à la représentation du sens interviennent secondairement pour bien montrer la faisabilité jusqu'à la représentation. Sous l'étiquette vague de variation de sens, nous incluons tout phénomène d'affinement ou de glissement du sens lexical d'un terme prédicatif sous l'influence de ses arguments lorsque celui-ci se trouve dans une proposition ou un énoncé. Nous incluons aussi des variations plus radicales telles que les métaphores et les métonymies. Le point de vue que nous défendons se situe à mi-chemin entre une approche de type lexicographique et une approche de type IA. En effet, l'approche énumérative, lexicographique, des différents emplois d'un terme paraît lourde, redondante et non structurée, et ne décrit pas les phénomènes au bon degré d'abstraction, handicaps importants pour des systèmes de TALN qui doivent tendre à une efficacité et à une robustesse maximale. Au contraire, l'approche de l'intelligence artificielle a tendance à trop généraliser en négligeant le langage et la dimension cognitive que nous voulons mettre en avant. Des représentations très génériques, peu adaptées aux traitements du TALN peuvent alors être proposées.

Nous tentons, expérimentalement, et sans prétendre que les choix et décisions soient uniques, de définir une voie moyenne qui développe une méthode (1) au niveau lexical, pour isoler des sens de lexèmes et (2) pour relier ces sens, lexicaux, aux usages rencontrés dans les propositions afin de déterminer le sens exact d'un prédicat en contexte. Nous avons aussi comme objectif d'aboutir à des descriptions les plus complètes possibles, permettant de simuler, par un modèle dynamique, les usages créatifs de la langue.

Dans ce cadre, les questions sont alors de déterminer quels sont les critères pour réaliser un compromis entre couverture linguistique, expressivité, efficacité, généricité et portabilité. La description à un bon niveau d'abstraction des phénomènes devra prévaloir afin de développer une certaine économie de moyens. La stabilité des descriptions sur des prédicats ayant des sens apparentés sera aussi un critère déterminant.

La problématique abordée étant très vaste, nous nous concentrerons, dans cet article, sur le verbe *couper*, un verbe courant, qui a fait l'objet de nombreuses controverses en anglais, et qui a un bon degré de polysémie. A partir d'une analyse rapide de propositions faites pour ce verbe, nous indiquerons la démarche que nous adoptons qui consiste en la hiérarchisation des sens de *couper*. A un premier niveau, une représentation en pôles sous spécifiés sera proposée. A des niveaux plus fins, des sens plus spécifiques pourront être atteints en spécialisant, modulant une représentation générique proposée pour le pôle principal.

2. Vers une méthode d'analyse des sens

2.1. Données et ressources utilisées

Nous avons tout d'abord extrait les occurrences du verbe *couper* du corpus du Monde de 1994. Concernant l'analyse de ce verbe, nous avons consulté les entrées de WordNet (Fellbaum 97) du verbe *cut*¹ pour une description de type base de données lexicale conçue pour le support électronique, le TLF pour une description lexicographique et un dictionnaire multi-lingue pour un point de vue plus conceptuel.

2.1.1. WordNet

WordNet propose 41 sens pour *cut* en distinguant par exemple le sens 1 (separate with or as if with an instrument, ex: *cut the rope*) du sens 29 (to reap, ex: *cut grain*), du sens 33 (to shorten, ex: *cut my hair*), du sens 38 (to reduce, ex: *cut a manuscript*), du sens 7 (to style, to tailor, ex: *to cut a dress*) et du sens 24 (function as a cutting instrument, ex: *this knife cuts well*).

Les critères de distinction de ces sens ne sont pas clairement explicités et sont assez disparates. Pour les sens 38 et 1, ce sont les conséquences et la célérité du procès qui ont été mises en valeur: si on coupe un texte, on le réduit et le raccourcit. Pour le sens 33, c'est la manière de couper qui entre en ligne de compte: en supprimant les extrémités (shorten as if by severing the edges or ends). Pour le sens 7, c'est une métonymie sur l'argument (l'objet final pour la matière dans laquelle est réalisé cet objet) qui est en jeu. Quant au sens 24, il s'agit d'une variante syntaxique du sens 1.

On a dans ce type d'approche une mise à plat qui fait perdre toute unité sémantique au mot. On est très en-deçà de ce que font habituellement les dictionnaires: aucune étiquette susceptible de relier certains sens entre eux, aucune condition d'emploi n'est, par exemple, présente. Les métonymies sur les arguments et les alternances syntaxiques sont présentées comme créant de nouveaux sens. De notre point de vue, il s'agit de phénomènes exprimables à un autre niveau et qui ne peuvent être confondus avec la polysémie des unités lexicales. Les emplois concernés peuvent en effet être considérés comme dérivant d'un sens initial.

2.1.2. TLF: Trésor de la Langue Française

Le TLF propose pour le verbe *couper* une distinction en deux rubriques, discriminées par l'intervention ou non d'un instrument tranchant. On retrouve un critère fréquemment utilisé dans les systèmes de TALN qui fonde la discrimination des sens des prédicats en fonction de la structure argumentale associée. On a ainsi la rubrique A. ('rompre un corps continu par l'intervention d'un instrument tranchant') comprenant la sous-rubrique A.2 ('diviser en plusieurs morceaux': *couper le pain*) et la rubrique B. ('sans instrument tranchant') avec la sous-rubrique B.1 ('partager matériellement un ensemble': *couper une droite, la route coupe la plaine*).

Il apparaît que la différence entre les exemples de A.2 et ceux de B.1 ne vient pas de la présence d'un instrument tranchant mais plutôt du niveau d'abstraction, de matérialisation de la séparation ou de la division. Le critère choisi n'est donc pas réellement pertinent.

Les observations précédentes sur ces différents types de ressources montrent que le découpage obtenu pour les sens d'un lexème, loin d'être universel, unique, dépend totalement des

1. A 90%, les usages de *cut* et de *couper* se recouvrent.

critères choisis. Nous allons donc présenter les critères qui nous semblent fondamentaux puis, à partir de ces critères, notre analyse en **pôles**² du verbe *couper*.

2.2. *Nos critères et notre analyse en pôles du verbe couper*

Les critères que nous avons retenus sont les suivants :

- le critère d'économie descriptive qui consiste à ne pas indiquer au niveau de l'unité lexicale ce qui peut être partagé par d'autres unités lexicales et donc énoncé à un autre niveau (comme les règles d'alternance, de métonymie sur les arguments...),
- le critère de structuration de l'information lexicale qui permet de factoriser un comportement conceptuel ou inférentiel plutôt que purement linguistique dans une représentation sous-spécifiée,
- le critère de faisabilité, c'est à dire la capacité à expliciter des mécanismes ayant un pouvoir explicatif, par exemple des règles génératives permettant de spécifier cette représentation sous-déterminée en tenant compte du cotexte et plus largement du contexte d'énonciation.

Pour la constitution du premier niveau de différenciation des pôles, nous classons les emplois en pôles selon le comportement conceptuel ou inférentiel qu'ils permettent à ce niveau. Des considérations sur les domaines ontologiques sur lesquels opèrent les prédicats sont aussi cruciaux. Des emplois relevant du même pôle devront alors partager ce comportement conceptuel 'de base' même si, à l'intérieur du pôle, ce comportement pourra être adapté, modulé et affiné suivant l'emploi. Par contre, des emplois relevant de pôles différents devront pouvoir se différencier par rapport à ce comportement initialement défini.

Un premier pôle d'attraction concerne les cas où *couper* est compris comme signifiant **l'arrêt d'un processus continu** regroupant les divers emplois: *couper le moteur, couper le chauffage, la parole, la conversation, le son, la radio, une émission, un film, la conversation, le gaz, l'électricité, l'eau, la lumière, le courant, le radiateur, l'aide, l'approvisionnement, les fonds, les crédits, les subventions, l'ambiance, la fièvre, la faim, l'appétit, l'envie, le souffle*

Un second pôle concerne les cas où *couper* est compris comme signifiant d'abord **une séparation physique** regroupant, par exemple: *couper le pain, le saucisson, un gâteau, la tête, le doigt, une robe, une droite, le blé, le champ,....*

Les emplois du pôle (1) de type 'couper X' partagent le fait qu'on pourra déduire que X ou quelque chose (un processus continu) lié à X s'est arrêté. Ils amènent à des inférences dans le domaine temporel.

Les inférences permises par les emplois du pôle (2) relèvent plutôt du domaine spatial ou physique avec l'idée de séparation déduite de la coupure.

Le pôle (2) doit ensuite être subdivisé en fonction de la nature de la séparation (réelle/ métaphorique), du degré de matérialisation de la coupure, du fait que l'objet est séparé en au moins

2. Nous préférons employer le terme pôle plutôt que le terme sens. Alors que ce dernier laisse entendre un découpage net et universel, le terme de pôle peut être interprété comme une 'tendance', un point d'attraction capable d'attirer à lui un ensemble d'usages.

2 morceaux distincts ou bien qu'il subit une altération n'entraînant pas de séparation totale (couper le doigt = érafler).

Par subdivisions successives, on obtient entre autres les sous-sens suivants pour le pôle (2):

1. emploi direct: séparer en plusieurs parties un objet homogène ou non (i.e. avec des parties identifiables): *couper un gâteau, la tête*.
2. blesser: *couper la joue*, une entaille est faite dans l'objet, il y a altération sans séparation.
3. isoler: *couper quelqu'un de ses amis*, la séparation peut être analysée comme une métaphore du domaine physique vers celui du psychologique ou de la communication.
4. raccourcir, diminuer: *couper un texte, une émission, un film (ambigu), un budget*, on ôte une partie, probablement non essentielle et le tout existe toujours. Ce sous-pôle est relié au pôle par l'idée d'amputation d'une partie.
5. traverser: séparation d'un espace, proche de 1. Ici, la coupure n'est pas réelle (ou reste très superficielle) mais a une dimension métaphorique interprétable par la perception: *la route coupe la forêt*.

Ce sens peut relater un état, via un glissement aspectuel, où le verbe de changement d'état perd son caractère agentif pour ne conserver que l'état résultant. C'est un glissement classique avec les verbes de mouvement, comme par exemple dans: *la route passe le col, la route longe la rivière...*

6. mordre: coupure métaphorique, proche de 2., liée au caractère vif ou incisif (métaphorique aussi) de l'objet: *le froid coupe la peau*, comme si le froid faisait des entailles dans la peau.

Comme on le voit, les instruments que l'on aurait pu utiliser pour réaliser l'action de couper sont peu présents comme critère discriminant. C'est davantage la nature sémantique des arguments ainsi que des éléments contextuels, inférentiels, qui sont importants pour distinguer ces sous-sens. La nature de la coupure, liée à l'objet, est elle aussi prépondérante.

On peut distinguer deux autres pôles, plus marginaux :

Un troisième pôle regroupe des emplois où *couper* signifie dénaturer, rompre une unité par introduction d'un élément d'une nature différente, avec *couper de l'eau avec du vin...*

Un quatrième pôle correspondant à un usage très spécifique regroupe les emplois avec la préposition à (*ne pas couper à Y*) et prend le sens d'éviter. On peut éventuellement considérer qu'il s'agit là d'un autre lexème puisque la préposition est solidaire du verbe (mais se déplace par 'pied piping').

Bien que ces pôles semblent a priori assez disjoints, on peut noter certaines continuités. Dans certains énoncés comme (a) et (b), on peut considérer que le premier et le second pôle se superposent.

(a) *couper les lignes d'approvisionnement des Serbes.*

(b) *L'inondation a coupé la route.*

Il y a en effet dans les deux cas une interprétation en pôle 2 avec une coupure spatiale, physique et une interprétation en pôle 1 correspondant pour (a) à l'arrêt de l'approvisionnement et pour (b) à l'arrêt du flux de voitures susceptibles de circuler sur la route³.

Dans d'autres cas comme dans *couper l'émission avec des pages de publicités, couper un voyage par une visite à des amis*, il y a à la fois du temporel : l'arrêt provisoire de X (*émission, voyage*) et du physique : on a intercalé physiquement des éléments de nature différente⁴.

D'autres énoncés peuvent être carrément ambigus : un pôle ou l'autre étant possible suivant le contexte dans lequel la phrase est prononcée ou suivant l'insertion de la phrase dans un texte plus large. Par exemple, le groupe verbal *couper un film* peut être attaché au pôle 2 dans le cas où on considère le film du point de vue 'objet-concret' (un ruban de pellicule) ; il s'agit alors du sens 'enlever au sens physique un morceau du film' comme dans (c). On peut aussi associer au film l'événement de sa projection et atteindre le pôle 1 comme dans (d).

(c) *Trop de longueurs: il faudrait couper le film au montage.*

(d) *Le machiniste a coupé le film avant la fin car il y avait une véritable émeute dans la salle.*

3. Éléments de modélisation

Ces pôles et ces sous-pôles étant définis, il faut d'une part les caractériser sémantiquement, définir les restrictions sur les arguments pour chacun d'eux, gérer les recouvrements et préciser les métaphores et métonymies qui peuvent s'appliquer. Il faut d'autre part, étant donné un usage (une phrase), se donner les moyens de déterminer le sens du prédicat, à partir des caractéristiques sémantiques de ses arguments. Les expériences que nous avons menées sur différents verbes ont montré que les restrictions de sélection sous la forme d'étiquette constante, empruntée à un treillis de types, aussi adapté soit-il, sont souvent insuffisantes voire inadaptées. Un travail en cours montre comment l'enrichissement des connaissances associées aux entrées lexicales peuvent être utilisées pour obtenir un ensemble de contraintes beaucoup plus adéquat. Ce point ne sera pas davantage développé ici, compte-tenu de sa complexité, des exemples seront donnés pour que le lecteur mesure les solutions à l'étude.

Au niveau de la modélisation, nous développons quelques représentations, dans le cadre de la LCS (Jackendoff 90, 97, Pinker 93), cadre qui nous semble bien organisé et conçu pour le problème que nous traitons. L'emploi de la LCS en TALN est davantage qu'un exercice de style: hors de la France, la LCS est souvent utilisée, en particulier comme forme 'inter-linguale' en traduction automatique.

Une représentation conceptuelle en LCS est composée de catégories sémantiques ('thing' regroupant 'human' et 'non-human', 'event', 'state', 'path', ...) indiquées en indice des crochets, de primitives conceptuelles (BE, GO, CAUSE, FROM, TO...) et de champs sémantiques qui indiquent ces primitives et qui permettent de préciser leur domaine d'application (localisation : +loc; propriétés : +char, +ident; psychologique : +psy...). Ainsi, la primitive $GO_{+char,+ident}$ indique un changement de propriétés, GO_{+loc} un changement de localisation.

3. Même si cette deuxième interprétation peut apparaître comme une conséquence de la première.

4. Ce sens de *couper* peut d'ailleurs être rapproché du pôle 3. La différence est qu'ici, il n'y a pas mélange entre les éléments de nature différente mais juxtaposition.

3.1. Le pôle 1: un modèle pour la notion de flux

La caractérisation du pôle 1 illustre parfaitement la difficulté à exprimer des restrictions de sélection en terme de catégorie sémantique préétablie : il n'existe pas une classe sémantique regroupant les divers arguments possibles *son, électricité, chauffage, crédits...* Par contre, la caractérisation peut se faire sur d'autres critères. En effet, on peut noter que ces différents lexèmes possèdent la capacité à évoquer un flux.

La représentation LCS de ce pôle dans le cas des constructions I (humain) coupe J peut être énoncée comme:

$$\left[\begin{array}{l} \text{[event CAUSE([thing I], [event GO}_{+char,+ident}([thing Flow(J)],} \\ \text{[path FROM}_{+char,+ident}([state Active]) TO}_{+char,+ident}([state NonActive])]])] \end{array} \right].$$

Active NonActive sont deux valeurs d'attributs de l'ontologie décrivant l'état d'un processus. On peut retrouver ce flux inférentiellement à partir de J, en exploitant des informations associées à J, par exemple, si l'on adopte le cadre du lexique génératif (Pustejovsky 91), celles pouvant être présentes dans le rôle télique de J.

Suivant la nature de J, ce flux peut être :

(a) **un flux d'argent** dans le cas des arguments de type *crédits, fonds, subventions, ...* avec pour télique, par exemple: *pay(X,Y,Z: argent)*⁵.

(b) **un flux d'informations** dans le cas de *parole, conversation, émission* et métonymiquement *radio, télévision, antenne, micro...*, le télique étant, par exemple: *say(X,Y,Z: information)*.

(c) **un flux d'énergie** pour *eau, gaz, électricité, moteur, lumière, robinet, ...*

(d) **un flux de ressource** *approvisionnements en armes, livraisons, aide alimentaire,...*;

Un flux peut être considéré, dans chaque cas, comme une arrivée récurrente 'd'objets de même nature' et c'est cette régularité temporelle qui est importante pour la combinaison avec *couper*.

Cette arrivée régulière peut se faire en suivant un support physique. Ainsi pour (c), le flux d'énergie parcourt des canalisations physiques ; dans ces cas-là, l'existence d'un moyen matériel de couper, sorte d'instrument installé sur le support physique, peut permettre des emplois métonymiques : le robinet, l'interrupteur...

La reconstitution du flux pour le moteur est sans doute moins directe. La télicité du moteur n'indiquera pas le flux ; par contre, on utilisera la connaissance qu'un moteur, pour réaliser sa tâche, doit être alimenté par un flux d'énergie qu'il transforme en énergie mécanique.

Pour faire apparaître le flux dans les emplois de type (a), on peut s'appuyer sur une analogie courante de l'argent vu comme un flux⁶. Pour autant, tous les lexèmes référant à de l'argent ne pourront s'employer avec *couper*. La destination de cet argent et l'aspect régulier du versement de cet argent sont importants. On peut considérer qu'aux noms en question sera associée une connaissance du type 'versé à X régulièrement tous les Y'.⁷

5. Comme pour la LCS, ces termes sont laissés en anglais pour éviter toute confusion entre niveaux de description.

6. l'argent circule, coule, passe de main en main, ...

7. A noter aussi que le pluriel joue un rôle pour cette interprétation.

Pour (b), le flux est clairement un flux sonore (ou d'images ou des deux), composé d'informations, facile à restituer. Dans le cas où ce flux est généré par un matériel audiovisuel, des métonymies du type 'objet émettant le flux pour le flux lui-même' (radio, télé) sont possibles et de la même façon que pour (c), des informations lexicales associées aux arguments seront exploitées.

Les arguments de type (d) peuvent être typés comme des procès dont la récurrence est soit intrinsèque, soit donnée par l'emploi du pluriel.

Pour les arguments de type *envie, faim, soif, appétit, désir...*, le flux est moins directement accessible ; il faut sans doute voir une métaphore de la notion de flux dans le domaine sensoriel où ces noms évoquent alors un processus continu correspondant à un manque éprouvé. *Couper*, dans ces cas-là, c'est mettre fin à cette sensation de manque.

La notion de flux reste à spécifier davantage de manière, notamment, à la distinguer de la notion de processus continu pour laquelle un verbe tel que *interrompre* pourrait être plus adapté. Traiter de façon précise des arguments acceptés par *couper* nécessite aussi d'examiner plus en profondeur les propriétés et les fonctions de ces flux afin d'aider à décider si un terme donné peut indiquer directement ou métonymiquement un flux et ainsi être un argument acceptable de *couper*.

3.2. Le pôle 2: dépendances entre représentations

Au niveau de ce pôle, nous montrons l'existence d'une représentation générique qui va se spécialiser pour chaque sous-pôle. Nous précisons aussi les restrictions de sélection associées.

(1) Couper = séparer physiquement, sens le plus immédiat: I coupe J, instrument et manière étant des modificateurs fréquents. I: animé, J: matériau ou objet naturel ou essentiellement naturel (bois, pain, métal), par contre peu d'objets complexes manufacturés peuvent être coupés. Il reste à définir une méthode d'identification et de caractérisation de ces objets, qui ne soit pas essentiellement lexicographique. La représentation LCS est la suivante:

$$\left[\begin{array}{l} \text{[event CAUSE}([\text{thing I}], [\text{event GO}_{+loc}([\text{thing J}], \\ \text{[path TO}_{+loc}([\text{state Discontinuous }])]])]) \end{array} \right]$$

Le terme 'Discontinuous' provient d'une ontologie décrivant l'état physique des objets⁸.

Les métonymies sont très nombreuses, entre autres celle où la totalité est utilisée pour une partie (s'agissant par exemple de l'instrument, le couteau pour la lame).

(2) Blessier I et J sont animés, J fait souvent l'objet de métonymies. L'instrument peut devenir sujet grammatical par alternance syntaxique, mais il ne fait pas partie de la structure argumentale *a priori*.

$$\left[\begin{array}{l} \text{[event CAUSE}([\text{thing I}], [\text{event GO}_{+char,+ident}([\text{thing J}], \\ \text{[path TO}_{+char,+ident}([\text{state Injured }])]])]) \end{array} \right]$$

On notera la similitude avec la représentation (1): il y a seulement changement d'ontologie dans la LCS: de la localisation (+loc) au domaine des propriétés (+char, +ident), et changement de la fonction ontologique élémentaire Discontinuous devenant Injured.

8. Ce terme est arbitraire et peut être modulé.

(3) **Séparer** I est un humain ou un événement (la maladie a coupé J de K), J est un humain (ou métaphoriquement un animal ou une plante) un groupe d’humains ou une structure sociale (l’humain pouvant être une métonymie des groupes ou vice-versa) et K est similaire à J. La métonymie ‘lieu pour structure sociale’ est fréquente.

$$[_{event} CAUSE([_{thing} I], [_{event} GO_{+psy\vecom}([_{thing} J], [_{path} AWAY - FROM_{+psy\vecom}([_{thing} K]])])])].$$

Nous sommes ici dans le domaine ontologique du psychologique (+psy) ou de la communication (+com), le chemin est un éloignement de K. Pour garder une plus grande similarité avec les deux cas ci-dessus, on pourrait avoir:

$[_{path} TO_{+psy\vecom}([_{state} Separated - from(K)])])]$. Mais l’idée de séparation fait partie de la sémantique de AWAY-FROM, quel que soit le domaine ontologique. Ici encore, le statut des fonctions ontologiques est à préciser dans un cadre plus global. Notons que la littérature abonde en ontologies nominales, mais qu’elle est très pauvre pour les ontologies des attributs des propriétés.

(4) **Raccourcir, diminuer** I: humain, J est un élément cognitif ou de type communication (un texte, un exposé, une émission, un discours, etc.) ou bien encore de type somme d’argent (budget, etc) ou des éléments plus spécifiques tels que herbe ou cheveux difficilement classables dans une même catégorie. Selon l’argument objet, ce ne sera pas la même dimension qui sera raccourcie ou diminuée: il faudra donc extraire dynamiquement de la structure lexicale de l’objet la dimension affectée.

$$[_{event} CAUSE([_{thing} I], [_{event} GO_{+char,+ident}([_{thing} J], [_{path} TOWARDS_{+char,+ident}([_{state} LESS(Prop)])])])].$$

Prop est une propriété prototypique de l’objet de type dimension, pour un texte: la longueur physique, pour un film: la durée, etc. Ces informations peuvent se trouver dans diverses structures selon l’approche. Dans notre cas, nous les trouverons dans le rôle constitutif, dûment typé, de la structure Qualia (Pustejovsky 91).

Le cas de ‘couper les cheveux’ peut se traiter ici, ou bien sous le cas (1), dans ce cas, le fait que les cheveux soient plus courts est une conséquence ‘inférée’ et non plus un élément sémantique lexical.

On se trouve ici à nouveau dans le domaine des propriétés. La primitive TOWARDS est utilisée au lieu de TO pour marquer le caractère progressif de l’opération.

Les autres cas se traitent de la même manière. Nous avons donné ici la représentation des verbes seuls. Celle de la proposition se construit compositionnellement à partir de la représentation du verbe et des arguments, modulo les restrictions de sélection. Les instanciations des champs sous-spécifiés se fait toujours de façon monotone croissante. Les primitives essentielles restent stables ainsi que la structure de la représentation (Saint-Dizier 98).

4. Conclusion, perspectives

Les aspects lexicaux et la polysémie en particulier sont la plupart du temps la pierre angulaire du TALN. Les travaux dans ce domaine deviennent de plus en plus approfondis, soit au niveau du contenu linguistique soit au niveau de leur lien avec les analyseurs ou les générateurs de

langage naturel. L'approche présentée se situe délibérément dans la perspective du contenu, qui est un préalable à tout traitement. Nous pensons qu'un retour sur une analyse en profondeur des données lexicales est nécessaire en particulier pour mieux exprimer les contraintes d'usages entre prédicat et argument(s), mieux modéliser la notion de sens, intégrer la prise en compte de la dimension dynamique du langage et la production de métaphores et de métonymies en particulier.

Même si le lien au TALN peut paraître un peu lointain, ce qu'il faut essentiellement retenir de ce travail, c'est la méthode et l'organisation des descriptions liées à diverses formes de raisonnement qui s'y rattachent. Il reste un travail de longue haleine pour définir et organiser par exemple les restrictions ainsi que pour valider cette approche sur d'autres classes de verbes.

Une comparaison des emplois prédicatifs faits en utilisant le verbe *couper* avec ceux employant les noms "d'action" correspondants *coupure*, *coupe*, *coupage* serait aussi fructueuse. Dans cet objectif, un travail utilisant des corpus spécialisés où les noms sont plus largement étudiés devrait être bientôt entrepris. Cela nous permettra aussi de confronter nos descriptions, faites a priori indépendamment d'une application ou d'une famille d'applications, à des occurrences attestées dans un domaine d'étude spécialisé.

Références

Fellbaum, C. (1997), A Semantic Network of English Verbs, in C. Fellbaum (ed) *Wordnet: An electronic Lexical Database*, Cambridge, MIT press.

Jackendoff, R., (1990), *Semantic Structures*, MIT Press.

Jackendoff, R., (1997), *The Architecture of the Language Faculty*, MIT Press.

Lakoff, G., Johnson, M. (1980), *Metaphors we Live By*, University of Chicago Press.

Levin, B., (1993), *English verb Classes and Alternations: A Preliminary Investigation*, Chicago Univ. Press.

Nunberg, G.D., Zaenen, A., (1992), Systematic Polysemy in Lexicology and Lexicography, proc Euralex92, Tampere, Finland.

Ostler, N., Atkins, S., (1992), Predictable Meaning Shifts: some lexical properties of lexical implication rules, in J. Pustejovsky and S. Bergler (eds.) *Lexical Semantics and Knowledge Representation*, Springer Verlag.

Pernelle, N., (1998), *Traitement automatique des polysémies relationnelles: utilisation et contrôle des règles d'extension de sens*, PhD dissertation, LIPN, Paris13.

Pinker, S., (1993), *Learnability and Cognition*, MIT Press.

Pustejovsky, J., (1991), The Generative Lexicon, *Computational Linguistics*, vol 17-4.

Pustejovsky, J., (1995), *The Generative Lexicon*, MIT Press.

Saint-Dizier, P., (1998), A Generative Lexicon Perspective for Adjectival Modification, in proc. ACL-Coling98, Montreal.

Saint-Dizier, P., (1999), Underspecification for Sense Variations, dans actes IWCS3, Tilburg, Pays-Bas.