Activation de l'information lexicale dans la combinaison adjectif-nom

Pablo Gamallo et Marie-Laure Reinberger

Laboratoire de Recherche sur le Langage
Université Blaise Pascal
34, avenue Carnot, 63000 Clermont-Ferrand
Tél: 04.73.40.63.37, Télécopie: 04.73.40.64.43
gamallo@lrl.univ-bpclermont.fr, reinberger@lrl.univ-bpclermont.fr

Résumé

La question de l'accès à l'information est présente dans tout travail portant sur l'organisation des structures lexicales et la caractérisation de mécanismes de combinaison de ces structures. La modélisation et le traitement de ce phénomène sont devenus des objectifs prioritaires des approches computationnelles de la sémantique. Cet article présente quelques éléments organisateurs de structures lexicales porteuses d'un certain degré d'indétermination. Il propose également quelques outils compositionnels qui régulent l'accès aux informations contenues dans les structures lexicales et qui engendrent le processus de construction dynamique de la signification. Ce travail a été effectué, ici, dans une perspective de simulation du processus d'interprétation d'expressions complexes de type "nom-adjectif".

1. Introduction

Cet article a pour principal objet l'étude de certains aspects de la structure lexicale des adjectifs et des noms. Nous partirons de l'hypothèse que les unités lexicales sont associées directement à l'ensemble des connaissances encyclopédiques. Dans cette approche non modulaire, on ne conçoit pas un module d'informations lexicales séparé d'un module d'informations non lexicales ou encyclopédiques. Il n'y a qu'un seul module d'informations, qui est, lui, mis en rapport d'une manière particulière avec le lexique (Langacker, 1991). L'un des phénomènes qui découle alors naturellement de la grande richesse encyclopédique rattachée aux unités lexicales est l'indétermination. Nous proposons ici, d'une part, la caractérisation d'éléments organisateurs des structures lexicales, éléments propres à appréhender des indéterminations, et, d'autre part, l'analyse d'outils de combinaison permettant de restreindre l'accès à cette richesse lexicale et donc les indéterminations lexicales. Ce travail est réalisé dans l'optique d'une simulation informatique du processus d'interprétation d'expressions complexes par un auditeur en compréhension. La mise en place d'objets et de mécanismes linguistiques s'accompagne donc d'une formalisation de ces différents éléments. En particulier, nous distinguons deux types d'éléments intervenant dans le processus d'interprétation des expressions de type nom-adjectif : les unités lexicales mises en relation, à savoir le lexème nominal et le lexème adjectival, et l'ensemble de mécanismes permettant la composition de ces deux lexèmes.

Nous présenterons ici des éléments permettant de structurer et d'organiser l'information encyclopédique associée aux lexèmes nominaux et adjectivaux, information qui est porteuse d'un certain degré d'indétermination. L'outil de modélisation que nous avons choisi d'utiliser pour la représentation des structures lexicales est le graphe. Ensuite, nous analyserons sommairement, et à titre d'illustration d'une mise en oeuvre des principes lexicaux, les mécanismes intervenant dans la combinatoire nom-adjectif. Pour ce faire, nous nous servirons d'un processus poly-opératoire dynamique que nous appelons “rôle-expert”. Ce processus déclenchera les mécanismes de co-activation de certaines facettes informationnelles des structures du nom et de l'adjectif, et de co-spécification informationnelle du nom par l'adjectif et de l'adjectif par le nom par création d'une nouvelle structure correspondant au sens porté par l'expression complexe.

2. Le lexique et l'accès à l'information

2.1. Les paquets conceptuels et leur notation en graphes

Un lexème est conçu comme un point d'accès à des informations conceptuelles de nature encyclopédique. Ces informations sont organisées autour de la notion de “paquet conceptuel” (Gamallo, 1998). Un paquet conceptuel est un objet informationnel défini à partir de la notion d'opération d'intégration. Les arguments de cette opération sont des unités informationnelles, plus ou moins complexes et plus ou moins spécifiques, qui sont mises en relation pour être intégrées dans une unité d'ordre supérieur, le résultat de l'opération. Le résultat de l'opération d'intégration est un paquet conceptuel intégrateur. Ce paquet est construit comme le “tout” par rapport aux arguments de l'opération, qui sont conçus, eux, comme ses paquets conceptuels internes, c'est-à-dire comme ses “parties constituantes”. En fait, un paquet conceptuel peut être pris comme argument de différentes opérations d'intégration et devenir une partie constituante de divers paquets intégrateurs. Un paquet conceptuel est donc doté d'une structure récurrente.

Un paquet conceptuel peut être caractérisé à partir des relations méréologiques “tout-partie” et “partie-tout” qu'il entretient avec d'autres paquets conceptuels. D'une part, un paquet conceptuel est relié à ses paquets internes, c'est-à-dire les éléments informatifs connectés par l'opération d'intégration pour construire le paquet conceptuel intégrateur. D'autre part, un paquet conceptuel est relié aux paquets intégrateurs pour lesquels il est un élément informatif constituant. Parmi les relations méréologiques, nous distinguons celles qui représentent des liens fonctionnels entre un participant et l'événement dans lequel il s'intègre (i.e. les rôles thématiques) de celles qui représentent des liens de constituance entre un ingrédient et l'entité complexe dont il fait partie (Winston *et al.*, 1987).

L'organisation de l'information en paquets conceptuels sera représentée sous forme de graphes. Nous nous référons pour cela à la définition des graphes conceptuels telle qu'on la trouve chez (Sowa, 1984), c'est-à-dire des graphes comportant deux types de nœuds : les nœuds-concepts et les nœuds-relations. Cependant, ici, nous emploierons le terme de “nœud” pour spécifier les nœuds-concepts, et celui de “relation” pour les nœuds-relations. Nous parlerons également de liens ou d'arcs associés à une relation. Un graphe conceptuel est un graphe fini et connexe. C'est un ensemble de nœuds reliés entre eux par des arcs étiquetés. Le schéma général d'un graphe est noté : [NŒUD] $\rightarrow$[NŒUD], où '$\rightarrow$' symbolise le lien associé à la relation qui unit ces deux nœuds. Une relation entre nœuds spécifie le rôle joué par ces nœuds dans le graphe.

Chaque nœud dispose d'un ou plusieurs liens, chacun devant être relié à un nœud. Voyons comment nous utilisons les graphes pour la représentation des paquets conceptuels et des relations méréologiques :
- un nœud représente un paquet conceptuel. Les paquets intégrateurs et les paquets internes mis en relation avec un paquet conceptuel seront décrits sous forme de nœuds reliés à ce paquet.
- une relation représente un lien méréologique entre paquets. Comme nous avons distingué deux types conceptuels de liens méréologiques, on distinguera les relations fonctionnelles, notées "$\Rightarrow$", des relations de constituance, notées "$\rightarrow$".

Soit le graphe :

$$\begin{array}{ll} \text{[phy]} & \Rightarrow[\text{év}_{dépl}] \\ & \rightarrow[\text{matière}] \end{array}$$

FIG. 1 – *Exemple de graphe*

Le nœud [phy] représente un paquet conceptuel caractérisant un objet physique. Dans ce graphe, il est relié méréologiquement à deux nœuds. D'une part, il est relié fonctionnellement au nœud [év$_{dépl}$], qui représente un paquet événementiel intégrateur. L'objet physique est donc pris ici comme une partie intégrée à l'événement, c'est-à-dire comme un de ses participants. Comme les liens ne sont pas étiquetés, le graphe ne précise pas le rôle fonctionnel particulier - agent, patient, ... - que l'objet [phy] joue dans l'événement. D'autre part, [phy] est relié par une relation de constituance à une de ses propriétés : le nœud [matière], représentant la matière de l'objet physique. L'objet physique est pris ici comme le tout intégrateur. La matière, quant à elle, est donc prise comme une des propriétés constituantes de l'objet. Un paquet conceptuel sera donc représenté par un noeud et relié au moyen de relations à ses paquets internes et à ses paquets intégrateurs, eux-mêmes représentés par des nœuds.

2.2. Paquets accessibles à partir du lexique

Un paquet conceptuel peut être porteur d'un contenu encyclopédique très riche. Comme il est relié aussi bien à ses paquets constituants qu'à ses paquets intégrateurs, il est un point d'accès à un réseau complexe d'informations. Du fait de cette richesse informationnelle, les paquets conceptuels ne peuvent pas être associés directement aux unités du lexique (les lexèmes). En fait, la structure du lexique impose des filtres restreignant l'accès à l'information encyclopédique portée par les paquets conceptuels. Prenons quelques exemples.

Supposons que, pour pouvoir rendre compte d'expressions comme 'voiture confortable', 'démarrer la voiture', on associe le lexème 'voiture' à un paquet conceptuel qui renvoie à un objet constitué, entre autres, de sièges et d'un moteur. Ces constituants - les sièges, le moteur, etc. -, en tant qu'entités matérielles, sont susceptibles de contenir des informations sur leur propre couleur, leurs propres dimensions, etc. Etant donnée la structure récurrente d'un paquet conceptuel, on devrait donc, à partir de ce paquet, accéder aux informations relatives à ces propriétés. Or, de telles informations sont difficilement accessibles à partir du lexème 'voiture'. La couleur et les dimensions des sièges et du moteur ne sont prises en compte, ni dans l'interprétation de 'voiture confortable' et 'démarrer la voiture', ni même dans l'interprétation d'expressions faisant référence directement à la couleur et aux dimensions de la voiture : 'voiture rouge' et 'petite voiture'.
Supposons que, pour pouvoir rendre compte des expressions 'lire un poème' et 'lire une feuille', on associe les lexèmes 'poème' et 'feuille' à un paquet représentant une configuration symbolico-

matériel, c'est-à-dire à une entité complexe caractérisée par une facette symbolique et par une facette physique sur laquelle le contenu symbolique se matérialise et se visualise. Etant donnée la structure récurrente d'un paquet conceptuel, on devrait accéder à partir de ce paquet particulier aux différentes informations concernant la facette symbolique et le support physique. Or, le lexème 'poème' ne donne pas accès à des informations propres au support physique : ??'poème bleu', ??'poème large' ; tandis que le lexème 'feuille' ne donne pas accès à des informations relatives exclusivement à la facette symbolique : ??'feuille naïve', ??'feuille dramatique'.
Supposons que, pour pouvoir rendre compte de certaines interprétations des expressions 'commencer le plat' et 'commencer le repas', on associe les lexèmes 'repas' et 'plat' à, entre autres, un paquet représentant l'action de manger quelque chose. Comme toute action perfective, elle renvoie à des informations concernant la durée de l'action. Or, tandis que le lexème 'repas' permet l'accès à cette information ('long repas'), le lexème 'plat' coupe un tel accès : ??'long plat'.

Ces exemples montrent qu'il est nécessaire de caractériser des principes régulant, à partir du lexique, l'accessibilité aux informations définies dans les paquets conceptuels. Nous définissons une unité lexicale comme un filtre qui, en fonction de certains principes, donne accès à un ensemble d'informations encyclopédiques. Les informations auxquelles il donne accès constituent ses "paquets accessibles". Les paquets accessibles à partir d'un lexème représentent en fait ses potentialités encyclopédiques. Plus un lexème possède des paquets accessibles, plus son potentiel encyclopédique est complexe et donc plus il peut être porteur d'un sens polysémique ou indéterminé. Nous ferons l'hypothèse que les liens méréologiques peuvent être caractérisés, soit comme des liens d'accessibilité à partir d'un lexème, soit comme des liens de non accessibilité à partir d'un lexème. Les liens d'accessibilité, non seulement relient deux paquets - un paquet de départ et un paquet d'arrivée -, mais présentent aussi le paquet de départ comme un point ouvrant l'accès aux informations du paquet d'arrivée. Les liens de non accessibilité, quant à eux, présentent les informations du paquet d'arrivée comme étant non accessibles depuis le paquet de départ. Dans la représentation en graphes, les relations fonctionnelles et de constituance ouvrant l'accès aux informations seront notées respectivement: "$\Rightarrow\succ$" et "$\rightarrow\succ$". L'accessibilité est donc notée par le symbole "$\succ$".

2.3. Paquets sources et schémas d'accès

Les paquets accessibles à partir d'un lexème sont reliés entre eux par des liens méréologiques et constituent un réseau d'informations. Dans ce réseau, on peut distinguer les paquets qui représentent les sources d'information du lexème : les "paquets sources". Un lexème est directement associé à ses paquets sources qui, à leur tour, sont reliés à tous les paquets accessibles du lexème. Les liens méréologiques entre un paquet source et un paquet accessible constituent un "schéma d'accès" informationnel à partir du lexème. Dans un schéma d'accès, les informations contenues dans les paquets accessibles peuvent être *transférées* au paquet source. Du fait qu'un paquet source est relié à un grand nombre de paquets accessibles, il représente alors le point de départ de divers schémas d'accès. De cette manière, un paquet source est le point d'accès à un faisceau d'informations, les informations contenues dans les schémas d'accès qui ont le paquet source comme élément de départ. Intuitivement, un paquet source est le point d'accès à un potentiel informationnel autour duquel peut être caractérisée une entité conceptuelle représentant un *référent* possible du lexème. Les noms, au moyen d'une détermination particulière, se caractérisent par le fait de construire dans le discours des entités référentielles. Les adjectifs, en revanche, ne sont pas associés directement à des référents discursifs. A la différence de la structure lexicale des lexèmes adjectivaux, celle des lexèmes nominaux sera organisée autour de paquets sources. C'est donc à partir de la notion de paquet source que nous essayerons de différencier ces deux types de structures lexicales. Dans la notation en graphes conceptuels, un paquet source d'un lexème est représenté comme un nœud d'entrée du graphe. Ce type de nœud sera noté en gras. Un nœud d'entrée est donc le point d'accès à un réseau d'informations associé au lexème. Prenons quelques exemples de structures lexicales.

2.3.1. Unités lexicales nominales

Les lexèmes 'planche', 'poteau', 'bâton', etc. désignent des objets physiques construits comme des configurations spatio-matérielles simples, au sens où elles ne sont pas composées d'autres objets internes. Du point de vue de la constituance, une configuration spatio-matérielle, notée [phy], se compose d'une matière occupant un espace (Borgo *et al.*, 1996). La matière, [matière], est la propriété de l'objet qui intègre des informations concernant la couleur, la texture, etc. L'espace borné, noté [espace], est la propriété de l'objet qui intègre les informations concernant sa forme et ses dimensions. Du point de vue de sa fonctionnalité, cette configuration est reliée à différents événement par des liens fonctionnels : l'objet physique peut jouer divers rôles fonctionnels, tels que "déplaçable", "touchable", etc. Les événements de déplacer et toucher sont à leur tour constitués de propriétés événementielles, telles que la vitesse ou le mode de réalisation. Dans le graphe de la figure 2, **[phy]** est le nœud de départ représentant le paquet source à partir d'un lexème comme, par exemple, 'bâton'. Les constituants internes de ce paquet et les événements physiques auxquels il est relié - par exemple des événements de déplacement : [év$_{dépl}$] - sont des paquets accessibles à partir de **[phy]**. Cependant, les propriétés internes de ces événements - [vitesse], [mode] - ne sont pas accessibles : ??'bâton rapide', ??'bâton tranquille', ??'bâton mouvementé', etc.

[phy]	$\rightarrow\succ$[matière]	$\rightarrow\succ$[couleur]
		$\rightarrow\succ$[texture]...
	$\rightarrow\succ$[espace]	$\rightarrow\succ$[dimensions]...
	$\Rightarrow\succ$[év$_{dépl}$]	$\rightarrow$[vitesse]
		$\rightarrow$[mode]

FIG. 2 – *Objet physique*

A partir du paquet source de cette structure, on peut caractériser plusieurs schémas d'accès. Par exemple : **[phy]** $\rightarrow\succ$[matière] $\rightarrow\succ$[couleur] représente un schéma d'accès.

Par ailleurs, un lexème nominal peut être associé à différents paquets sources. Chaque paquet source caractérise le lexème par rapport à une perspective ou point de vue référentiel. Le degré d'indétermination d'une unité lexicale est directement lié à la diversité des paquets sources qu'on lui associe. Prenons quelques exemples.
Les lexèmes 'pomme', 'livre', 'cahier', ... sont associés, au moins partiellement, à une structure d'informations représentant une configuration spatio-matérielle complexe. Cette configuration est constituée par une entité physique externe et par une entité physique interne. La structure lexicale de ces unités se caractérise par le fait que, aussi bien la configuration complexe que l'entité interne peuvent être conçues comme des paquets sources : une pomme avec ou sans peau est encore une pomme, un livre avec ou sans couverture est encore un livre. Cependant, l'entité externe (la peau de la pomme ou la couverture du livre) ne sont pas pris comme des points d'entrée à partir du lexème. Le graphe de la figure 3 représente un aspect de la structure lexicale d'un lexème comme 'pomme'. Les deux paquets sources de cette structure sont la configuration complexe, notée **[phy_config]**, et l'entité physique interne, notée **[phy_int]**. Notons que la couleur et

la texture de l'entité interne ne sont pas des paquets accessibles, car c'est la couleur et la texture de la peau qui déterminent la couleur et la texture de la pomme : 'pomme rouge', 'pomme lisse'. Ces propriétés peuvent être transférées à configuration complexe **[phy_config]**. Par contre, une action de type *toucher*, notée [év$_{toucher}$], est accessible à partir de l'entité externe et de l'entité interne : l'interprétation de 'toucher une pomme' implique que l'on touche soit la peau externe, soit la partie interne. Cette fonctionnalité peut donc être transférée aux deux paquets sources : **[phy_config]** et **[phy_int]**. Il y a d'autres fonctionnalités, par exemple l'événement particulier [év$_{éplucher}$], qui concernent directement la configuration complexe : l'interprétation de 'éplucher une pomme' met en jeu ensemble la peau externe avec la facette interne.

[phy_config]	$\rightarrow\succ$[phy_ext]	$\rightarrow\succ$[matière]	$\rightarrow\succ$[couleur]
			$\rightarrow\succ$[texture]
			$\rightarrow$[goût]
		$\rightarrow\succ$[espace]	$\rightarrow\succ$[dimensions]...
		$\Rightarrow\succ$[év$_{toucher}$]	
	$\rightarrow\succ$**[phy_int]**	$\rightarrow\succ$[matière]	$\rightarrow$[couleur]
			$\rightarrow$[texture]
			$\rightarrow\succ$[goût]
		$\rightarrow\succ$[espace]	$\rightarrow\succ$[dimensions]...
		$\Rightarrow\succ$[év$_{toucher}$]	
	$\Rightarrow\succ$[év$_{éplucher}$]		

FIG. 3 – *Pomme*

A partir du paquet source **[phy_config]** de cette structure, on peut définir différents schémas d'accès. Par exemple : **[phy_config]** $\rightarrow\succ$**[phy_int]** $\rightarrow\succ$[matière] $\rightarrow\succ$[goût] est un schéma d'accès. Attention, dans ce schéma d'accès le paquet **[phy_int]** n'est pas pris comme le paquet source point de départ. Mais à partir du paquet source **[phy_int]** de cette structure, on peut définir différents schémas d'accès. Par exemple : **[phy_int]** $\Rightarrow\succ$[év$_{toucher}$] est un schéma d'accès.

Les structures lexicales de 'livre', 'poème' et 'feuille' partagent un même contenu conceptuel, mais elles sont organisées autour de différents paquets sources et donc de différents schémas d'accès. Le lexème 'livre' peut être associé à, au moins, trois paquets sources (voir graphe de la figure 4). Premièrement, on a un paquet représentant une configuration spatio-matérielle, notée **[phy]**. Nous ne prenons pas en compte ici la complexité de cette configuration qui, dans le cas d'un livre, a une structure similaire à celle représentée dans le graphe de la figure 3. Cette configuration spatio-matérielle est le point de départ des schémas d'accès à des informations sur des propriétés et des fonctionnalités purement matérielles : 'livre bleu', 'toucher le livre'. Deuxièmement, on a un autre paquet représentant une entité symbolique, et notée **[symb]**. Cette entité est le point de départ des schémas d'accès à des propriétés et fonctionnalités purement symboliques : 'livre naïf', 'critiquer un livre'. Enfin, on a un paquet intégrant les deux paquets précédents, et noté **[phy_symb]** (en référence au type pointé de (Pustejovsky, 1995)). Ce paquet intégrateur est caractérisé, d'une part, par les informations transférées à partir des deux paquets constituants et, d'autre part, par le fait qu'il donne un accès direct à des fonctionnalités qui mettent en jeu ensemble les facettes spatio-matérielle et symbolique : 'lire/écrire un livre'.

Par opposition à 'livre', les lexèmes 'poème' et 'feuille' sont associés uniquement à deux paquets sources. Le graphe de la structure lexicale de 'poème'i est organisé autour de l'entité symbolique **[symb]** et de la configuration complexe physico-symbolique **[phy_symb]**. Les informations portées par la facette physique ne sont pas accessibles et donc ne peuvent pas

$$
\begin{array}{lll}
\textbf{[phy\_symb]} & \rightarrow\succ\textbf{[phy]} & \rightarrow\succ\text{[matière]} \\
 & & \rightarrow\succ\text{[espace]} \\
 & & \Rightarrow\succ[\text{év}_{toucher}] \\
 & \rightarrow\succ\textbf{[symb]} & \rightarrow\succ\text{[forme]} \\
 & & \rightarrow\succ\text{[contenu]} \\
 & & \Rightarrow\succ[\text{év}_{critiquer}] \\
 & \Rightarrow\succ[\text{év}_{lire/écrire}] &
\end{array}
$$

FIG. 4 – *Objet physico-symbolique de type livre*

être transférées à la configuration complexe : ??'toucher le poème', ??'poème bleu'. La structure lexicale de 'feuille' présente une organisation "inverse" de celle de 'poème' : les paquets sources sont **[phy]** et **[phy_symb]**, et les informations non accessibles sont celles portées par la facette symbolique : ??'feuille naïve', ??'critiquer la feuille'.

La structure lexicale de 'repas', quant à elle, est organisée autour de deux paquets sources (graphe de la figure 5). Premièrement, on a un paquet représentant une substance matérielle, **[mat_subs]**, donnant accès, d'une part, à des informations concernant la quantité et la composition - 'repas copieux', 'repas végétarien' -, d'autre part, à des informations fonctionnelles concernant les événements de manger et de préparer le repas : $[\text{accomp}_{préparer}]$ et $[\textbf{accomp}_{manger}]$. Les événements notés [accomp] représentent des accomplissements, c'est-à-dire des événements perfectifs ayant une durée de réalisation. Deuxièmement, l'événement même de manger, $[\textbf{accomp}_{manger}]$, est aussi pris comme un paquet source. Cet événement donne accès à toutes ses propriétés constituantes : le mode ('repas mouvementé'), la durée ('repas rapide', mais également 'pendant le repas', 'long repas', etc.). Notons que, par rapport au paquet **[mat_subs]**, la durée de la préparation est aussi un paquet accessible, car l'interprétation de 'repas rapide' non seulement peut concerner l'événement de manger le repas, mais aussi sa préparation. Un tel accès est cependant partiel puisque les expressions 'long repas' ou 'pendant le repas' ne concernent que la consommation du repas.

$$
\begin{array}{lll}
\textbf{[mat\_subs]} & \rightarrow\succ\text{[quantité]} & \\
 & \rightarrow\succ\text{[composition]} & \\
 & \Rightarrow\succ[\textbf{accomp}_{manger}] & \rightarrow\succ\text{[durée]} \\
 & & \rightarrow\succ\text{[mode]} \\
 & \Rightarrow\succ[\text{accomp}_{préparer}] & \rightarrow\succ\text{[durée]} \\
 & & \rightarrow\text{[mode]}
\end{array}
$$

FIG. 5 – *Repas*

L'ensemble de schémas d'accès caractérisés autour des différents paquets sources associés à une unité lexicale représente son potentiel encyclopédique, potentiel qui est, lui, porteur d'un certain degré d'indétermination.

2.3.2. Unités lexicales adjectivales

A la différence des unités lexicales nominales, la structure conceptuelle d'un lexème adjectival n'est pas organisée autour de paquets sources. Dans la combinaison entre un adjectif épithète et un nom, ce sont précisément des paquets sources du nom qui serviront de point de départ pour organiser l'information de la structure conceptuelle associée à l'expression complexe. Dans une telle combinaison, les informations associées à l'adjectif sont prises fondamentalement comme des contraintes conceptuelles qui restreignent les schémas d'accès du nom susceptibles d'être utilisés dans la structure conceptuelle de l'expression complexe.

A titre d'exemple, nous allons représenter divers aspects de l'organisation conceptuelle de deux adjectifs : 'rouge' et 'rapide'. Les représentations ne seront que partielles, dans la mesure où elles doivent permettre d'illustrer le phénomène de combinaison adjectif-nom. La structure lexicale de 'rouge' est représentée dans le graphe de la figure 6. Cette structure est organisée autour des paquets contenant les informations spécifiques véhiculées par le lexème : [couleur_rouge] et [contenu_communiste]. Ces paquets sont représentés comme étant accessibles à partir d'autres paquets qui peuvent faire partie de différents schémas d'accès du nom. Dans cette structure, nous distinguons, au moins, trois types de schémas d'accès possibles. Premièrement, quand l'adjectif se combine avec un nom désignant une configuration spatio-matérielle ou une substance matérielle, on construit un schéma d'accès vers le paquet spécifique [couleur_rouge] à partir de [matière] et [phy] : 'encre rouge', 'feuille rouge'. Notons que dans le cas de 'encre rouge', le paquet source du schéma est [matière], car l'encre est conçue comme une substance matérielle et non comme une configuration spatio-matérielle. Deuxièmement, quand l'adjectif se combine avec un nom désignant une entité symbolique, c'est à partir du paquet [symb] que l'on construit le schéma d'accès vers [contenu_communiste] : 'poème rouge'. Troisièmement, quand l'adjectif se combine avec un nom désignant une configuration physico-symbolique, c'est à partir du paquet [phy_symb] que l'on construit le schéma d'accès vers les deux paquets spécifiques [couleur_rouge] et [contenu_communiste] : 'livre rouge'. Notons que, comme 'livre' peut aussi désigner un objet physique [phy] et une entité symbolique [symb], l'interprétation de l'expression 'livre rouge' reste indéterminée. Elle peut donner lieu à, au moins, trois lectures : un objet physique de couleur rouge, une entité symbolique d'idéologie communiste ou un objet de couleur rouge ayant un contenu symbolique d'idéologie communiste.

$$\begin{array}{lllllll}
\text{[couleur_rouge]} \prec\!\leftarrow & \text{[matière]} & \prec\!\leftarrow & \text{[phy]} & \prec\!\leftarrow & \text{[phy_symb]} \\
\text{[contenu_communiste]} & & \prec\!\leftarrow & \text{[symb]} & \prec\!\leftarrow &
\end{array}$$

FIG. 6 – *Rouge*

La structure lexicale de l'adjectif 'rapide' est représentée dans le graphe de la figure 7. Cette structure est organisée autour des paquets contenant les informations spécifiques véhiculées par le lexème : [vitesse_rapide] et [durée_courte]. Ces paquets sont accessibles à partir d'autres paquets qui peuvent faire partie d'au moins deux schémas d'accès différents. Premièrement, on peut construire un schéma d'accès vers le paquet [vitesse_rapide] à partir du paquet [act], qui représente une activité, c'est-à-dire un événement imperfectif : 'mouvement rapide', 'dactylo rapide'. Notons que dans le cas de 'mouvement rapide', le paquet [act] est le paquet source, car le mouvement est associé directement à un événement imperfectif, et non à une entité physique comme 'dactylo'. Deuxièmement, on peut aussi construire un schéma d'accès vers [durée_courte] à partir du paquet [accomp] : 'repas rapide'. Notons que dans le sens d'une consommation rapide, c'est le paquet [accomp] qui représente le paquet de départ du schéma, tandis que dans le sens d'une préparation rapide, c'est la substance matérielle, [mat_subs], qui devient le paquet de départ ou paquet source.

$$\begin{array}{lll}
\text{[vitesse_rapide]} & \prec\!\leftarrow & \text{[act]} \\
\text{[durée_courte]} & \prec\!\leftarrow & \text{[accomp]}
\end{array}$$

FIG. 7 – *Rapide*

3. Le rôle-expert modifieur

3.1. Présentation

Le processus de combinaison entre plusieurs lexèmes est mis en oeuvre par un rôle-expert. C'est lui qui, grâce aux structures mises en place pour la représentation des lexèmes, va élaborer une construction qui permette un choix parmi plusieurs interprétations possibles de l'expression traitée. Il y a un rôle-expert pour chaque type de combinaison catégorielle. Leur appel est géré par un gestionnaire central. Le rôle-expert se présente sous la forme d'un processus composé d'une suite d'instructions (en référence aux "mots-experts" de (Small & Rieger, 1982)). Il est lancé après lecture de la combinaison lui correspondant. Il agit de manière dynamique sur les différents composants de l'expression tout au long de la construction d'une représentation de la signification. On peut diviser l'action du rôle-expert en deux types de tâches, qui se déroulent cependant de manière conjointe :
- une tâche de co-activation, car chaque élément influera sur la détermination des schémas d'accès à activer dans la structure des autres éléments.
- une tâche de co-spécification informationnelle ou intégration conceptuelle, qui consiste en la création d'une nouvelle unité conceptuelle constituée des schémas d'accès activés dans les différents graphes.

Notons que l'efficacité du rôle-expert est rendue possible par la représentation des structures lexicales que nous avons construites. Notons également qu'ici, le rôle expert est lancé sur une combinaison syntaxique déterminée et qu'il agit différemment selon la nature catégorielle des composants de l'expression.
Ici, le rôle-expert dont il sera question est celui dévolu à la combinaison d'un nom avec un adjectif épithète, c'est le rôle-modifieur. Le rôle-modifieur sera donc lancé d'emblée après la lecture d'un nom et d'un adjectif épithète. Les structures lexicales que nous avons établies traduisent le fait que le nom est caractérisé comme un faisceau de schémas d'accès organisés autour de paquets sources, alors que l'adjectif se présente comme un ensemble de contraintes pouvant être utilisées pour sélectionner des schémas d'accès. En conséquence, le rôle-modifieur n'agira pas de la même manière sur le nom et sur l'adjectif.

Le rôle-modifieur ne combine pas tout le potentiel informationnel accessible à partir des deux lexèmes. Il agit sur le graphe de l'adjectif en déterminant la signification de l'adjectif selon le contexte créé par le nom, et sur le nom en restreignant la quantité d'information accessible en fonction des éléments apportés par l'adjectif. On est donc en présence du phénomène interactif dynamique de co-activation décrit plus haut, et qui s'accompagne d'une co-spécification génératrice d'une unité de représentation de l'interprétation de l'expression. Du point de vue de la modélisation, le rôle-modifieur effectue un parcours des graphes conceptuels et établit un lien de transmission d'information s'il y a adéquation entre un schéma d'accès du graphe du nom et un chemin dans le graphe de l'adjectif. Si le rôle active plusieurs schémas d'accès du nom et plusieurs paquets de l'adjectif, l'interprétation de l'expression complexe est indéterminée et il sera créé entre les deux structures autant de liens qu'il y a de schémas activés. L'interprétation de l'expression complexe sera alors traduite par une unité conceptuelle plus élaborée et plus ou moins riche conceptuellement selon le nombre de schémas activés, qui correspond au paquet conceptuel de l'expression. En résumé, l'accès à l'information encyclopédique a bien été restreinte, mais plusieurs interprétations restent possibles. Par la suite, cette expression peut être déterminée dans un nouveau processus de combinaison et par un nouvel appel au processus rôle-modifieur.

3.2. Exemples de combinaisons de deux lexèmes par le rôle-modifieur

Nous avons choisi de traiter deux exemples mettant en jeu des lexèmes nominaux présentant plusieurs paquets sources pour lesquels la combinaison avec certains adjectifs laisse subsister une indétermination. Traitons pour commencer le cas de la combinaison de 'rouge' avec un lexème représentant une configuration physico-symbolique comme 'livre'. Le graphe de représentation de 'livre' porte l'information selon laquelle c'est la couleur de la couverture, autrement dit le constituant externe, qui est pertinente lorsque la couleur du livre est spécifiée. Il contient en conséquence un schéma d'accès jusqu'au paquet représentant la couleur de la couverture du livre. La présence de ce schéma d'accès va permettre au rôle-modifieur d'établir une correspondance entre la propriété de couleur de la couverture et la propriété de couleur d'un objet physique de 'rouge', puis la création d'un lien permettant un transfert d'information entre ces deux éléments.
De la même manière, un lien sera créé reliant le paquet spécifiant le contenu d'un livre avec la facette spécifiant un contenu idéologique communiste associé à 'rouge'. On a au final la création de trois chemins d'accès, en considérant également celui qui spécifie un livre rassemblant les deux caractéristiques. On est donc en présence d'une indétermination qui sera peut-être levée dans le déroulement du discours par l'arrivée d'informations complémentaires, pouvant porter par exemple sur la visualisation du livre (et privilégier à ce moment le schéma d'accès spécifiant la couleur de la couverture), ou sur un jugement impliquant un contenu d'idéologie d'extrême-gauche et pointant donc sur le schéma mettant en relief le contenu symbolique de l'ouvrage. Dans ce cas, le rôle-modifieur est rappelé avec ces nouvelles informations et est susceptible de modifier la structure en désactivant certains schémas d'accès.
Le graphe de la figure 8 est celui résultant de la construction de la combinaison 'livre rouge'. Il utilise les structures exposées dans les graphes des figures 3 et 4 pour une configuration complexe et un objet physico-symbolique. [phy_ext] représente la couverture du livre, et [phy_int] l'ensemble des pages. Dans les graphes, le lien de transmission d'information créé lors de la co-spécification par le rôle-modifieur est noté $\Longleftrightarrow$. Un nœud activé est représenté en majuscules. Les représentations de structures lexicales ne sont pas complètes. Nous n'avons représenté que les éléments pertinents pour l'interprétation de l'expression.

				[COULEUR_ROUGE]
				⇕
[PHY_SYMB]	→≻**[PHY_CONFIG]**	→≻[PHY_EXT]	→≻[MATIERE]	→≻[COULEUR]
				→≻[texture]
			→≻[espace]	→≻[dimensions]
		→≻[phy_int]	→[matière]	→[couleur]
				→[texture]
			→≻[espace]	→≻[dimensions]
	→≻**[SYMB]**	→≻[forme]		
		→≻[CONTENU]		
		⇕		
		[CONTENU_COMMUNISTE]		

FIG. 8 – *Livre rouge*

Voyons maintenant la combinaison 'repas rapide'. Dans le cas de 'repas rapide', on a accès à deux significations qui entraînent donc une indétermination. Il peut s'agir de la consommation du repas ou de sa préparation. Le rôle-modifieur identifiera une première fois le paquet [durée_courte] du lexème 'rapide' avec le paquet du lexème 'repas' spécifiant la durée de l'accomplissement consistant à manger le repas, et une deuxième fois avec le paquet du lexème 'repas' spécifiant la durée de la préparation du repas. Ici aussi, des informations ultérieures permettront peut-être de restreindre les possibilités d'indétermination, provoquant alors un rappel du rôle-modifieur et une modification par celui-ci de la structure existante.

[MAT_SUBS]	$\rightarrow\succ$[quantité]		
	$\rightarrow\succ$[composition]...		
	$\Rightarrow\succ$**[ACCOMP$_{manger}$]**	$\rightarrow\succ$[DUREE]	$\Longleftrightarrow$[DUREE_COURTE]
	$\rightarrow\succ$[mode]		
	$\Rightarrow\succ$[ACCOMP$_{préparer}$]	$\rightarrow\succ$[DUREE]	$\Longleftrightarrow$[DUREE_COURTE]
		$\rightarrow$[mode]	

FIG. 9 – *Repas rapide*

Dans le cas de 'repas rapide', comme dans celui de 'livre rouge', le rôle-modifieur procède à un marquage de l'expression comme indéterminée, afin que le gestionnaire des rôles puisse transmettre cette information aux rôles-experts qui se dérouleront dans la suite du traitement.

4. Conclusion

Nous avons exposé dans cet article des principes d'organisation du lexique, et en particulier les différences de structuration entre les lexèmes nominaux et les lexèmes adjectivaux relativement à ces principes. Nous avons également tenté de caractériser les mécanismes internes à la fonction de modifieur. Pour ce faire, nous avons mis en relief la façon dont le rôle-modifieur agit de manière dynamique à l'intérieur même des structures lexicales pour activer les éléments pertinents dans la construction d'une interprétation. On pourrait à présent envisager d'affiner les notions relatives aux principes d'organisation et aux mécanismes combinatoires que nous avons présentés, ceci en spécifiant d'autres types syntaxiques de lexèmes qui mettront en jeu d'autres rôles-experts.

Références

BORGO S., GUARINO N. & MASOLO C. (1996). Stratified ontologies: the case of physical objects. In *Workshop on Ontological Engineering, ECAI-96*.

GAMALLO P. (1998). *Construction conceptuelle d'expressions complexes : traitement de la combinaison nom-adjectif*. Clermont-Ferrand: Thèse de Doctorat, Université Blaise Pascal.

LANGACKER R. (1991). *An introduction to Cognitive Grammar*, volume 1,2. Stanford: Standford University Press.

PUSTEJOVSKY J. (1995). *The Generative Lexicon*. Cambridge: MIT Press.

SMALL S. & RIEGER C. (1982). Parsing and comprehending with word experts (a theory and its realization). In W. LEHNERT & M. RINGLE, Eds., *Strategies for Natural Language Processing*, p. 89–148.

SOWA J. (1984). *Conceptual Structures: Information Processing in Mind and Machine*. London: Addison-Wesley.

WINSTON M., CHAFFIN E. & HERRMANN D. (1987). A taxonomy of part-whole relations. *Cognitive Science*, **11**(4), 417–444.