L'adjectif en français et à travers les langues, Presses Universitaires de Caen, 2004
Version préliminaire

Polysémie adjectivale et rection nominale : quand *gros* et *gras* sont synonymes

Jean-Luc Manguin, Jacques François, Bernard Victorri

Le propos de cet article est de présenter une méthode de reconnaissance de l'aptitude d'un régissant nominal à réduire à plus ou moins radicalement la polysémie d'un adjectif dépendant et plus spécialement à repérer la zone de recouvrement entre les champs sémasiologiques[1] de deux adjectifs aux emplois multiples, *gros* et *gras*, en couplant deux outils d'ingénierie linguistique,

(i) d'une part le DICTIONNAIRE ELECTRONIQUE DES SYNONYMES [abrégé DES] élaboré en 1997 par Sabine Ploux et Bernard Victorri (cf. Ploux & Victorri 1998) et actuellement accessible sous deux variantes légèrement différentes sur le site Internet du CRISCO et sur celui de l'Institut des Sciences Cognitives de Lyon

(ii) d'autre part la partie de la base de données textuelles FRANTEXT dont les textes ont été catégorisés par l'INaLF (environ 450 œuvres du 19° et 20° siècle).

Dans la première section nous donnerons une première image superficielle de la polysémie des deux adjectifs telle qu'elle ressort de l'inventaire des synonymes délivré par le DES et nous évoquerons le mode de constitution de cet inventaire. Nous montrerons ensuite comment la représentation visuelle de la polysémie lexicale et plus précisément adjectivale dans un espace multidimensionnel réalisée à partir du calcul de la distance entre 'cliques' de synonymes permet d'associer à des zones de synonymie des types de représentations contextuelles (Section 2). Puis nous exposerons la méthode de reconnaissance des collocations adjectivo-nominales, c'est-à-dire des co-occurrences de fréquence élevée dans la base FRANTEXT catégorisée qui permet d'identifier des contextes nominaux propres à favoriser la synonymie (Section 3) et en conclusion nous évoquerons les améliorations techniques auxquelles nous sommes en train de procéder tant en ce qui concerne le mode de visualisation du champ sémasiologique[2] que dans la collecte des co-occurrences adjectivo-nominales ainsi que l'évaluation de ce mode de représentation comme simulation de l'organisation du lexique mental de locuteurs francophones de diverses tranches d'âge mise en évidence par une série d'expériences psycholinguistiques menées en collaboration avec Françoise Cordier au laboratoire *Langage & Cognition* de l'Université de Poitiers.

---

1 En sémantique lexicale, deux types de champs sémantiques sont traditionnellement distingués : les champs onomasiologiques qui rassemblent des unités lexicales dénotant des objets, propriétés ou procès conceptuellement proches et les champs sémasiologiques qui sont une mise en ordre de la variété des types d'emploi des lexèmes polysémiques.

2 Voir la contribution de Bernard Victorri dans ce volume.

1. *Gros* et *gras* sont occasionnellement synonymes

En présence d'unités lexicales polysémiques — c'est-à-dire de la quasi-totalité du lexique d'une langue, en dehors du domaine, vaste mais cependant limité, des termes de langue de spécialité n'ayant (pas encore) donné lieu à des emplois métaphoriques — la pratique des dictionnaires de langue et des dictionnaires explicatifs des synonymes est de 'discrétiser' la polysémie lexicale dans la micro-structure de l'article lexicographique à l'aide d'une définition qui tire parti de différents outils, essentiellement le relevé de co-occurrences régulières, la substitution de synonymes et d'antonymes propres à conserver ou inverser le sens du terme vedette dans le type de contexte ainsi identifié et la limitation à un registre stylistique et/ou à un domaine de spécialité particuliers.

L'inventaire des relations synonymiques du DES a été constitué à partir des listes de renvois synonymiques de sept dictionnaires compilées par l'INaLF pour l'élaboration du *Trésor de la Langue Française*. Chaque renvoi n'a été comptabilisé qu'une fois et symétrisé, c'est-à-dire que le renvoi inverse a été introduit s'il était absent de cette compilation. Cette méthode rappelle le travail effectué il y a un quart de siècle par André Kahlmann (cf. Kahlmann A. (1975)). La base est constituée de deux dictionnaires de langues extensifs, le *Grand Larousse de la Langue Française* (1971) et le grand *Dictionnaire Robert* (1992)[3], de quatre dictionnaires de synonymes explicatifs, deux du 19e siècle, celui de Guizot (1864) et celui de Lafaye (1878) et deux du 20e, le dictionnaire de Bénac (1956) et celui de Bailly (1969) et d'un dictionnaire de synonymes non explicatif, celui de Du Chazaud (1989). L'inventaire ainsi constitué ne garde pas le souvenir de la discrétisation opérée par ces auteurs (à l'exception de Du Chazaud), c'est-à-dire que les relations de synonymie sont décontextualisées. Nous allons nous arrêter dans un premier temps à cet inventaire brut avant de passer à l'exploitation des cliques de synonymes. Le DES fournit une liste alphabétique de 147 synonymes de *gros* :

> *abondant, adipeux, ample, appuyé, arrondi, assourdissant, ballonné, barrique, beau, beaucoup, bedon, bedonnant, biquet, bombé, bon, bouffi, boulot, boursouflé, caniculaire, charnu, cher, colossal, commun, conséquent, considérable, copie, coquet, corpulent, courtaud, développé, dodu, élevé, éminent, empâté, en bloc, enceinte, enflé, énorme, épais, épanoui, essentiel, étouffant, excessif, excrément, exorbité, extraordinaire, fabuleux, faramineux, fécond, fertile, formidable, fort, gênant, gidouillard, globalement, globuleux, gonflé, grand,* ***gras****, grave, gravos, gros bonnet, grossi, grossier, haut, huile, hypertrophié, immense, important, imposant, influent, insupportable, intense, joli, joufflu, large, légume, lourd, maous, masse, massif, mastar, mastoc, mastodonte, membru, monstrueux, nabab, nombreux, notabilité, notable, nourri, obèse, obscène, opulent, ordinaire, pansu, paquet, patapouf, pataud, patoche, pénible, pépère, pesant, piffre, plein, porc, pote, potelé, poupard, poussah, principal, proéminent, profond, puissant, ragot, rebondi, renflé, replet, résonnant, retentissant, riche, robuste, rond, rondelet, rudimentaire, sérieux, sévère,*

---

[3] **Indiquer les éditions réellement utilisées.**

simple, solide, sonore, soufflé, spacieux, supérieur, tonneau, trapu, trivial, tuméfié, vaste, ventripotent, ventru, vif, vilain, violent, volumineux, vulgaire

et une liste plus courte de 71 synonymes de *gras* :

abondant, adipeux, amiteux, ample, arrondi, beau, bien en chair, bouffi, boulot, boursouflé, charnu, charnu, corpulent, dégoûtant, dodu, égrillard, empâté, épais, farfelu, fécond, fertile, fort, gaulois, généreux, glissant, gluant, graisse, graisseux, grasset, grassouillet, graveleux, grivois, ***gros****, grossier, huileux, lard, leste, libre, licencieux, malsonnant, moelleux, obèse, obscène, onctueux, opulent, ordurier, pansu, pâteux, plantureux, plein, poisseux, polisson, porc, potelé, poupard, productif, rabelaisien, rebondi, replet, répugnant, riche, rond, rondelet, rondouillard, sale, salé, suiffeux, suintant, vaste, ventru, visqueux*

Ces listes ne sont à proprement parler exploitables que pour les locuteurs aptes à s'imaginer des contextes d'emploi où la relation synonymique est pertinente.

— On note d'abord pour *gros* que certains emplois sont exclusivement nominaux, désignant soit des objets {*un gros*[-humain] ≈ *une barrique*, *une copie, un excrément*, *un tonneau*, etc.} soit des êtres humains {*un gros*[+humain] ≈ *un nabab*, *une notabilité*, *un notable,* etc.}. Cela vaut à un moindre titre pour *gras* {*le gras* ≈ *la graisse, le lard*}.

— Ensuite on remarque que certains synonymes de *gros* ou *gras* ne le sont que dans des contextes extrêmement restreints. Ainsi *caniculaire* ne peut se substituer à *gros* que comme qualification d'une chaleur ou d'une température. Dans le même domaine météorologique, un *gros temps* sera inversement un temps de forte dépression.

— En cherchant à s'imaginer des contextes nominaux, on s'aperçoit également que la possibilité de co-occurrence avec un même régissant nominal ne garantit pas la substituabilité. Un pré peut être salé et/ou gras, mais un *pré salé* est rarement *gras* et vice versa. Cependant *salé* est bien synonyme de *gras* comme qualification d'une plaisanterie ou d'une histoire et entre dans un sous-champ sans évaluation dépréciative constitué par {*gras, leste, salé, égrillard, gaulois, grivois, licencieux, polisson, rabelaisien*} aux confins d'un autre sous-champ constitué de termes dépréciatifs comme *graveleux*, *obscène* ou *ordurier*. Comparons par exemple (cf. Tableau 1) les conditions de synonymie entre {*gras, abondant, bouffi, corpulent, épais, graisseux, grivois, salé*} dans les contextes nominaux {*papier, herbe, personne, plaisanterie*} :

N *gras(se)*	***herbe***	***papier***	***personne***	***plaisanterie***
abondant	SYN	≠	*	*
bouffi	*	*	SYN	*
corpulent	*	*	SYN	*
épais	SYN	≠	±SYN	±SYN
graisseux	≠	SYN	≠	*
grivois	*	*	≠	SYN
salé	≠	?	*	SYN

Tableau 1 : Co-occurrence synonymique (SYN), non synonymique (≠) et non co-occurrence (*) entre quatre noms et huit synonymes de *gras*

Par exemple, *épais* se révèle synonyme de *gras* pour une herbe comme pour une plaisanterie, mais une plaisanterie ne peut pas être abondante alors qu'une herbe abondante est une herbe grasse, et en outre une plaisanterie grasse peut également être qualifiée de grivoise ou de salée alors qu'une herbe ne peut pas être grivoise et qu'une herbe salée n'est pas nécessairement grasse. Dans le premier cas, la raison du blocage est que les paires *herbe abondante* et *herbe épaisse* ne sont équivalentes à *herbe grasse* que comme désignations d'un espace herbeux et non d'un brin d'herbe particulier. L'image commune est celle de brins d'herbes qui poussent en masse serrée, or une plaisanterie est une production langagière délimitée et non collective, l'image ne peut pas y être transférée. Mais la plaisanterie peut être presque indifféremment qualifiée de *grasse* ou d'*épaisse* si elle manque de finesse, parce qu'appliqués à des objets physiques ou des personnes, *gras* et *épais* ont un même antonyme, *fin*. En revanche un papier fin s'oppose à un papier épais, mais non à un papier gras, parce que ce caractère ne peut pas lui être conféré par ses propriétés propres mais par le contact avec de la graisse, d'où la synonymie entre *gras* et *graisseux* que permet *papier* et qu'il ne partage avec aucun des trois autres noms. Enfin les dictionnaires nous apprennent qu'une plaisanterie grasse est une plaisanterie qu'on tenait typiquement dans les jours gras qui précédaient le carême. Il s'agit donc ici d'une métonymie, de la qualification des jours vers celle des comportements typiques de ce moment de l'année. En dehors de ces jours de licence culturellement encadrée, une plaisanterie grasse est difficile à faire passer, comme l'est un plat épicé, ce qui permet de comprendre qu'une plaisanterie grasse, par son caractère licencieux, soit aussi une plaisanterie salée.

— Inversement le relevé de tous les contextes [N *gras/grasse(s)*] dans la base FRANTEXT catégorisée révèle que certains noms qui régissent *gras* ne peuvent régir en relation synonymique aucun des synonymes de la liste, ainsi *foie gras* (45 occurrences), *mardi gras* (27), *plante grasse* (12, *matière grasse* (5), *crayon gras* (2).

— Le fait que *gras* apparaisse dans la liste des synonymes de *gros* et vice-versa laisse supposer que les champs sémasiologiques des deux adjectifs ont un espace commun. En comparant les deux listes, on constate effectivement qu'ils ont 33 synonymes en commun, soit 23% des synonymes de *gros* et 46% de ceux de *gras* (cf. Figure 1)

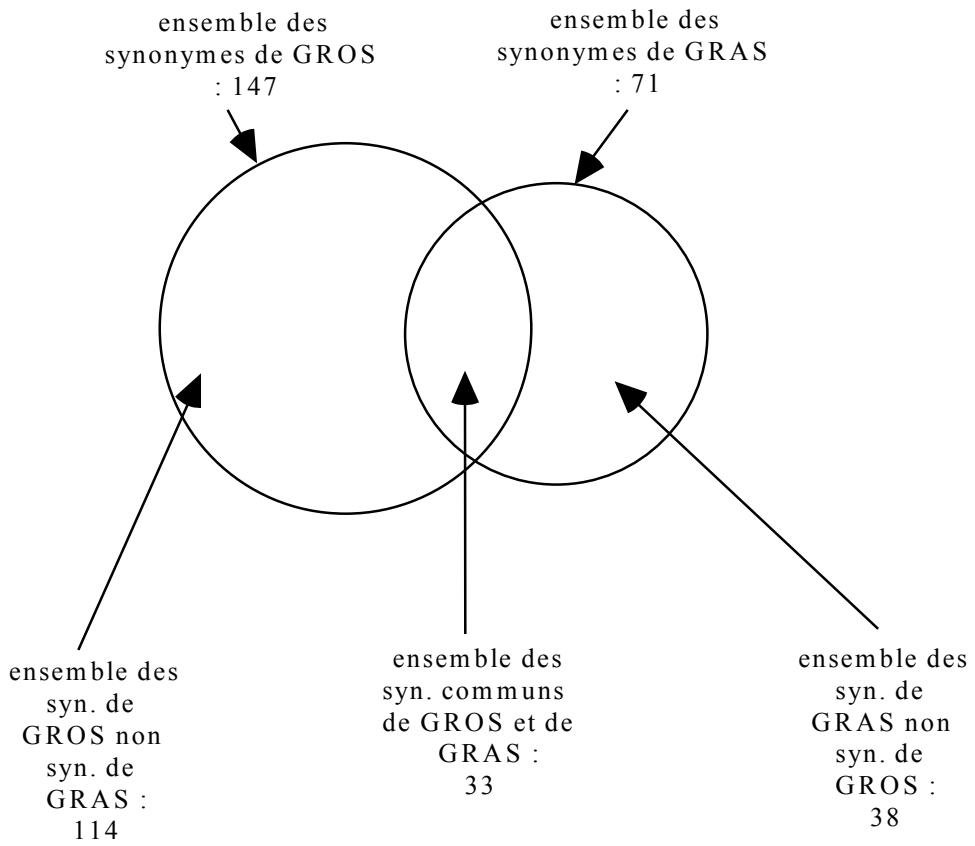


Figure 1. L'espace sémantique commun aux deux listes des synonymes de *gros* et *gras* fournies par le DES

Mais parler d'un espace sémantique commun laisse entendre que cet espace soit unifié. Or force est de constater que si la plupart des adjectifs de cet espace peuvent qualifier des personnes ou des parties du corps où le caractère *gros* est supposé en rapport avec un excès de graisse, d'autres s'appliquent à des comportements ou des actes {*grossier, obscène*} et *abondant* désigne l'effectif élevé d'un collectif (*herbe*, *cheveu*)[4]. Cet espace commun est donc lui-même composite. Seule l'analyse des regroupements des relations synonymiques en 'cliques' permet de clarifier la composition interne des trois espaces, la partie du champ sémasiologique de *gros* qui lui est propre, celle qu'il partage avec le champ sémasiologique de *gras* et la partie de ce dernier que *gras* ne partage pas avec *gros*.

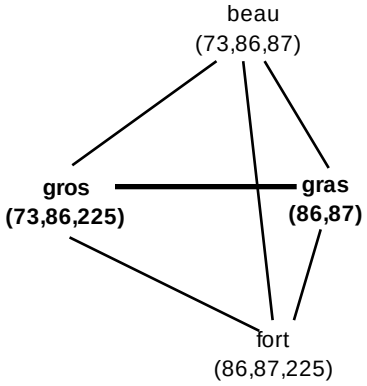
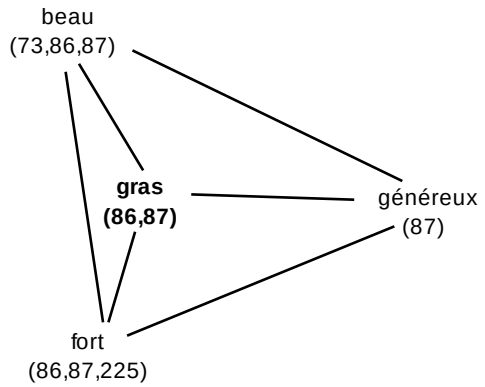
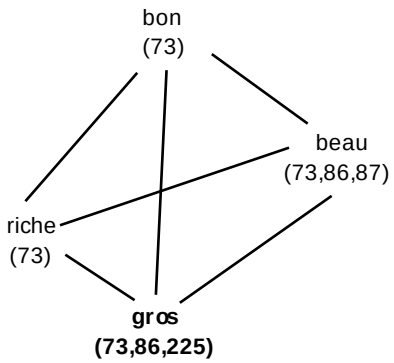
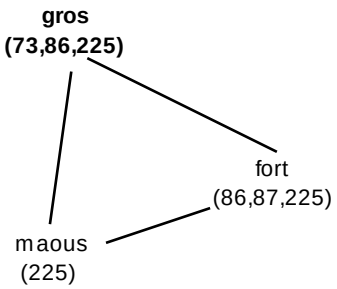
2. La représentation de l'espace sémantique multidimensionnel de *gros* et de *gras*

Nous appliquons ici une méthode qui a déjà été exposée par ailleurs (cf. Ploux & Victorri 1998), aussi nous nous contenterons d'en rappeler les principes. Tout d'abord, nous regroupons les synonymes de *gros* et de *gras* dans un même ensemble. Puis, nous recherchons les relations de synonymie existant au sein de cet ensemble ; nous obtenons ainsi un graphe, dans lequel nous recherchons toutes les cliques maximales ; l'expérience nous a cependant montré [Manguin J.L.& Victorri B. (1999)] qu'il vaut mieux ne retenir que les cliques contenant au moins un des mots-vedettes de départ (en l'occurrence *gros* ou *gras*). Au sein d'une clique, tous les éléments sont en relation (de synonymie) les uns avec les autres, mais cela ne les empêche pas de former d'autres cliques avec d'autres sommets du graphe. On obtient ainsi souvent une sorte de "tuilage" entre les cliques.

Afin de clarifier le phénomène de 'tuilage' entre les cliques, nous pouvons nous concentrer sur quatre cliques de faible effectif (trois à quatre constituants), telles que l'une d'entre elles, la 86 inclut *gros* et *gras* et les trois autres (87, 73 et 225) soit *gros*, soit *gras* et l'un au moins des autres constituants de la 86 (cf. Figure 2). La clique 86 rassemble les liens synonymiques entre *gros*, *gras*, *beau* et *fort. Gras*, *beau* et *fort* réapparaissent dans la 87 qui leur adjoint le lien synonymique avec *généreux*. Aucune autre clique ne comporte de lien entre *gros* et *généreux*, il n'y a donc pas de clique associant *généreux* au matériau constitutif de la clique 86. La 73 associe *riche* et *bon* à *gros* et *beau*, co-présents dans la 86. Enfin la 225 associe *maous* à *gros* et *fort* . La représentation combinée des quatre cliques (qui ne constituent évidemment qu'un minuscule sous-ensemble des 315 cliques de *gros* et/ou *gras*) met en évidence le 'tuilage' entre ces cliques, la clique 86 qui occupe une position centrale en raison de la co-présence de *gros* et *gras* apparaissant en traits gras (sans jeu de mots !). On notera qu'aucune des 315 cliques ne présente de lien synonymique entre *bon* et *généreux*, adjectifs

[4] *Herbe* et *cheveu* ne peuvent être *abondants* qu'en tant que collectifs.

pourtant manifestement synonymes. Cela est dû au fait que *gros* ou *gras* joue un rôle de filtre, c'est-à-dire qu'on ne trouve pas de triplet {*bon gros/gras, généreux*}.

<table>
<tr>
<td>Clique 86 {beau, fort, gras, gros}

</td>
<td>Clique 87 {beau, fort, généreux, gras}

</td>
</tr>
<tr>
<td>Clique 73 {beau, bon, gros, riche}

</td>
<td>Clique 225 {fort, gros, maous}

</td>
</tr>
</table>

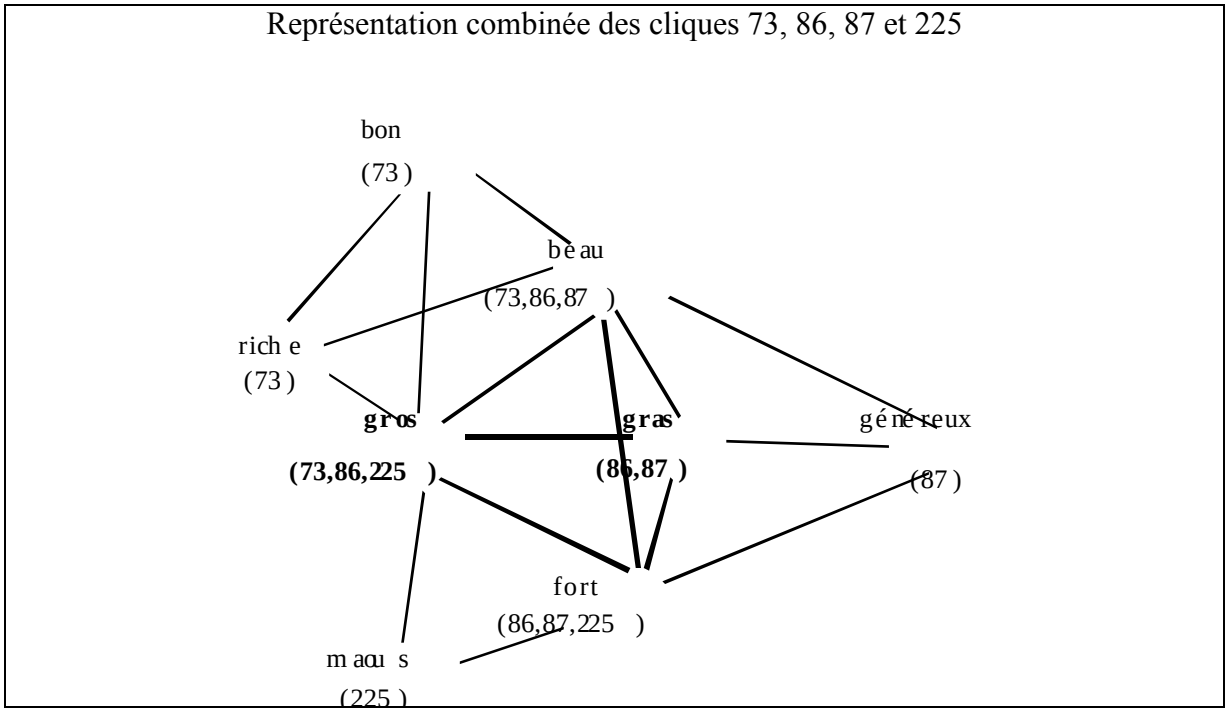


Figure 2. Représentation éclatée et combinée de quatre cliques

Pour passer à l'étape suivante, qui est la représentation de l'espace sémantique des mots-vedettes, il faut maintenant assimiler chaque clique à un vecteur d'un espace multidimensionnel (précisément, sa dimension est égale au nombre des synonymes) ; comme chaque axe correspond à un synonyme, la coordonnée de chaque clique sur cet axe sera égale à un ou zéro, suivant que la clique contient ou non ce synonyme.

Il s'avère que pour une représentation correcte, il est nécessaire de tenir compte de la fréquence d'apparition de chaque synonyme dans les cliques, ainsi que du nombre de synonymes dans chaque clique ; les coordonnées des cliques sont ainsi pondérées, et l'on va effectuer sur l'ensemble de ces données une Analyse en Composantes Principales, qui permet de dégager les plans d'observation les plus pertinents. Dans ce nouvel espace, les cliques sont toujours des points, et sur les figures planes données ici, les points sont les projections des points de l'espace sur le plan de la figure.

Chaque synonyme est représenté par un contour qui forme l'enveloppe convexe des points (donc des cliques) qui le contiennent. Il faut toutefois signaler que dans le cas qui nous occupe, la lecture de la représentation complète nécessite de consulter les différentes projections du nuage de points dans plusieurs plans principaux ; après cette phase, il est possible de combiner les plans les plus intéressants et d'en obtenir un nouveau qui offre une vue plus synthétique ; nous avons déjà employé cette opération dans des études antérieures [Manguin J.L.& Victorri B. (1999)] . La lecture d'une telle représentation devient alors plus aisée.

On voit ainsi, sur la figure, une zone commune à *gros* et à *gras*, et deux zones exclusives qui correspondent à chacun de ces adjectifs. Nous pouvons ainsi dire que dans ce plan, ce qui est *gras* est situé à gauche et ce qui est *gros* à droite.

Afin de favoriser la lecture des visualisations, nous nous concentrerons dans la figure 3 sur l'espace des synonymes de *gras* (et accessoirement de *gros*) et dans la figure 4 sur celui des synonymes de *gros* (et accessoirement de *gras*). Les deux figures 3 et 4 ont en commun la représentation de l'espace des deux adjectifs *gras* (à gauche) et *gros* (à droite). Sur la figure 3 on a représenté (en parcourant la représentation de haut en bas) l'espace des adjectifs *fécond*, *graisseux*, *dégoûtant*, *gaulois* et *potelé*. L'une des trois cliques jalonnant l'espace de *fécond* inclut *gros* , {*abondant ; fertile ;* ***fécond*** *;* ***gras*** *;* ***gros***}. Il en est de même pour l'une des cliques de *potelé*, {*arrondi ;* ***gras*** *;* ***gros*** *;* ***potelé*** *; rebondi ; rond*}. En revanche aucune des cliques constitutives de l'espace de *dégoûtant*, *graisseux* ou *gaulois* n'inclut *gros*[5]. En haut (espace de *fécond* / *fertile*) les synonymes spécifient une valeur positive de *gras* concernant la productivité (d'une terre, d'une prairie, d'une herbe). Au centre la valeur dépréciative véhiculée par *dégoûtant*, *répugnant*, *sale*, *graisseux* l'emporte. En bas à gauche (espace de *gaulois*) c'est l'idée de licence (donc de liberté excessive mais tolérée) qui prévaut avec *licencieux*, *grivois*, *égrillard*. En bas à droite (espace de *potelé*) on trouve la propriété physique d'une forme ronde (cf. *arrondi*, *rebondi*, *rond*) associée à un développement de la graisse dans une partie du corps, l'adjectif étant plus rarement appliqué à l'ensemble du corps d'un humain (*il est potelé*, ? *rond*, ?? *rebondi*, ?? *arrondi*).

[5] La représentation adoptée donne l'illusion que certaines des cliques de *dégoûtant* comportent *gros*. En fait, comme le montre la clique qui constitue l'extrémité droite de l'espace de *dégoûtant* {***dégoûtant*** *;* ***gras*** *; grossier ; obscène ; répugnant*} *gros* en est exclu.

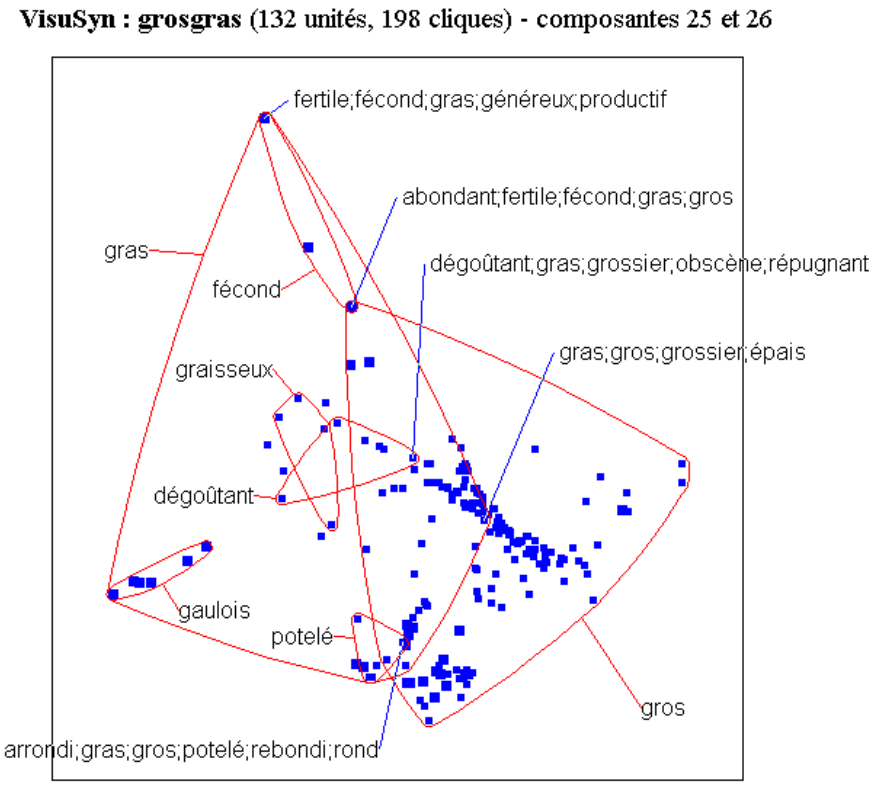


Figure 3 : espace sémantique de *gros*, *gras*, 3 synonymes de *gras* et 2 synonymes communs aux deux.

Sur la figure 4 consacrée aux synonymes de *gros*, nous avons mis en évidence les espaces respectifs de deux adjectifs, *abondant* et *adipeux* , qui empiètent sur celui de *gras*. Mais ces deux espaces sont disjoints et de fait ils incluent un type différent de caractère gras. Pour *abondant* (voir la clique {***abondant*** *; fertile ; fécond ;* ***gras*** *;* ***gros***}) c'est à nouveau la productivité qui est en cause, tandis que pour *adipeux* (***adipeux*** *;* ***gras*** *; grassouillet ; rebondi ; replet ; rondouillard*) c'est l'excès de graisse dans l'ensemble du corps d'un humain (et non d'une partie du corps, ce qui rapproche *adipeux* de *corpulent* et l'éloigne de *potelé*). L'espace de *gros* sur cette représentation a un centre spécifié par l'espace de *colossal* (caractérisant une taille excessive dans toutes les dimensions), une zone haute où la grosseur est synonyme de fécondité, une zone basse où elle est limitée à un gonflement soit de longue durée et qualifiant un humain dans sa globalité (espace de *ventru* qui recouvre celui de *bedonnant, ventripoten*t, etc.) soit transitoire et qualifiant une partie du corps (espace de *tuméfié* associé à *ballonné*, *boursouflé*, *enflé*, etc.) et enfin une zone droite marquée par l'espace de *vulgaire* (voir la clique {*commun ;* ***gros*** *; grossier ; ordinaire ; trivial ;* ***vulgaire***}) et celui très spécifique de *pénible* (*grave ;* ***gros*** *; lourd ; pesant ;* ***pénible***).
On observe que, contrairement à ce qu'on pourrait supposer, *vulgaire*, *trivial*, *ordinaire* n'apparaissent que comme synonymes de *gros* et *grivois*, *égrillard*, *licencieux* seulement comme synonymes de *gras*. Cela tient sans doute à une connotation négative dans le champ de *gros* (l'absence de finesse, de classe exprimée par *grossier*, *ordinaire* ou *commun*), tandis

que la licence bénéficie d'une (coupable) bienveillance dans le champ de *gras* (toujours en rapport avec la nécessité de la décompression sociale pendant la "semaine grasse" précédant le Carême, cf. §1)

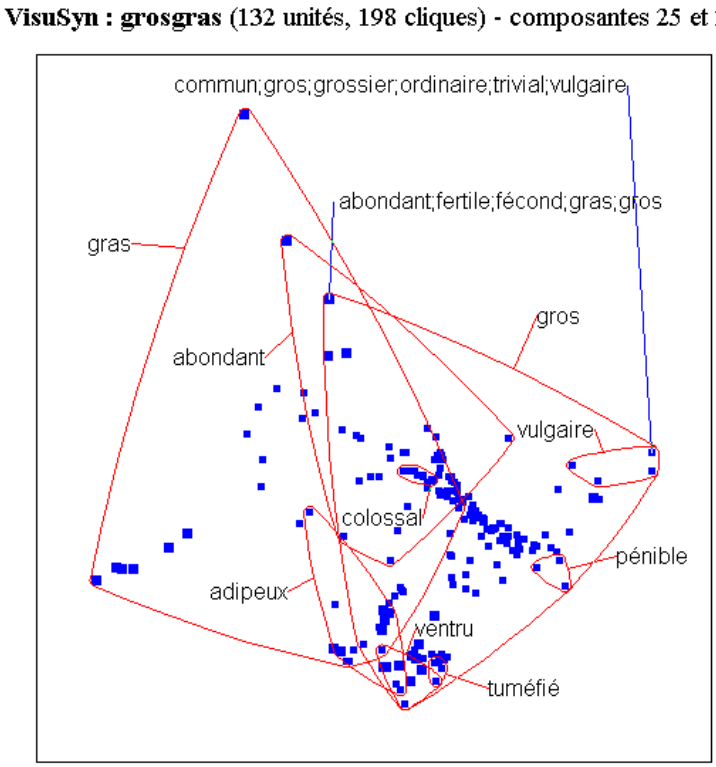


Figure 4 : espace sémantique de *gros*, *gras*, 4 synonymes de *gros* et 2 synonymes communs aux deux.

3. L'utilisation de la base Frantext catégorisée et l'ajout du contexte dans la représentation précédente

Pour ajouter à notre représentation les noms qui co-occurrent avec les adjectifs étudiés, nous avons choisi d'utiliser le corpus Frantext catégorisé, car il nous permet de rechercher des éléments en fonction de leur catégorie grammaticale (ce qui est impossible dans la base Frantext non catégorisée). Notre but est d'obtenir une liste de paires (Nom, Adjectif), où chaque Adjectif est un synonyme de *gros* ou *gras*, et où chaque Nom co-occurre au moins une fois avec *gros* ou *gras*.

Pour obtenir cette liste, nous avons procédé en deux étapes :

1) Tout d'abord nous cherchons les noms en cooccurrence avec *gros* ou *gras*, par deux requêtes Frantext données ci-dessous :

&e(g=S) &?(&e(g=Adv)) &e(c=(&mgras | &mgros) g=A) (A)

&?(&e(g=Adv)) &e(c=(&mgras | &mgros) g=A) &e(g=S) (B)

Ces requêtes se traduisent ainsi :

A: un substantif + un adverbe éventuel + un adjectif, forme fléchie de *gros* ou *gras*

B: un adverbe éventuel + un adjectif, forme fléchie de *gros* ou *gras* + un substantif

L'analyse des réponses de Frantext par programme nous donne une liste de noms (que nous appellerons *liste_noms*).

2) Une requête sur notre dictionnaire des synonymes nous avait déjà donné une liste d'adjectifs, que nous appellerons *liste_adjectifs*.

La deuxième étape consiste donc à effectuer les recherches suivantes dans Frantext :

&e(g=&lliste_noms) &?(&e(g=Adv)) &e(c=&lliste_adjectifs g=A) (A)

&?(&e(g=Adv)) &e(c=&lliste_adjectifs g=A) &e(g=&lliste_noms) (B)

Il faut noter que les deux listes constituées pour cette demande lors de la session Frantext contiennent non pas les mots simples, mais leurs formes fléchies.

Nous avons recueilli par cette méthode quelques 20 000 co-occurrences qui se regroupent en 1842 paires (nom, adjectif) ; le Tableau 2 donne les paires obtenues avec les trois noms *affaire*, *plaisanterie* et *ventre* par ordre de fréquence décroissante. Ces trois exemples ont été choisis parce qu'ils correspondent à des zones différentes sur la représentation.

Reg. Nom.	Adj.	Nb occ.
	grand	51
	gros	48
	important	35
	sérieux	34
	grave	21
	sale	16
	simple	14
	considérable	10
	colossal	5
	vilain	5
	pénible	4
affaire	*énorme*	3
	principal	3
	vaste	3
	immense	2
	lourd	2
	extraordinaire	1
	formidable	1
	haut	1
	large	1
	ordinaire	1
	retentissant	1
	supérieur	1

Reg. Nom.	Adj.	Nb occ.
	énorme	7
	lourd	6
	obscène	5
	grossier	4
	gras	3
	vulgaire	2
	colossal	1
plaisanterie	*dégoûtant*	1
	égrillard	1
	grand	1
	massif	1
	ordinaire	1
	sale	1
	salé	1

Reg. Nom.	Adj.	Nb occ.
	gros	69
	énorme	23
	ballonné	13
	gonflé	10
	rond	9
	enflé	6
	large	6
	arrondi	5
	bombé	4
	lourd	4
	rebondi	3
	rondelet	3
	obscène	2
ventre	*puissant*	2
	empâté	1
	épanoui	1
	fécond	1
	haut	1
	immense	1
	important	1
	monstrueux	1
	poisseux	1
	renflé	1
	solide	1
	spacieux	1
	trapu	1
	visqueux	1

Tableau 2 : Co-occcurrences rencontrées dans la base de texte entre les substantifs *affaire*, *plaisanterie*, *ventre*, les adjectifs *gras* et *gros* et leurs synonymes répertoriés par le D.E.S.

L'ensemble des paires observées et de leurs nombres d'occurrences respectifs va ensuite alimenter le programme de visualisation précédemment décrit ; par une technique déjà utilisée [François J., Victorri B., Manguin J.L.(à paraître)], nous calculons le degré de co-occurrence de chaque nom avec chacune des cliques du graphe de la synonymie, puis la "fonction de compatibilité" du nom avec les différents sens présents sur la représentation (Figure 5). Pour des raisons de visibilité, il n'est pas possible de faire figurer sur le graphique les fonctions de plusieurs noms simultanément.

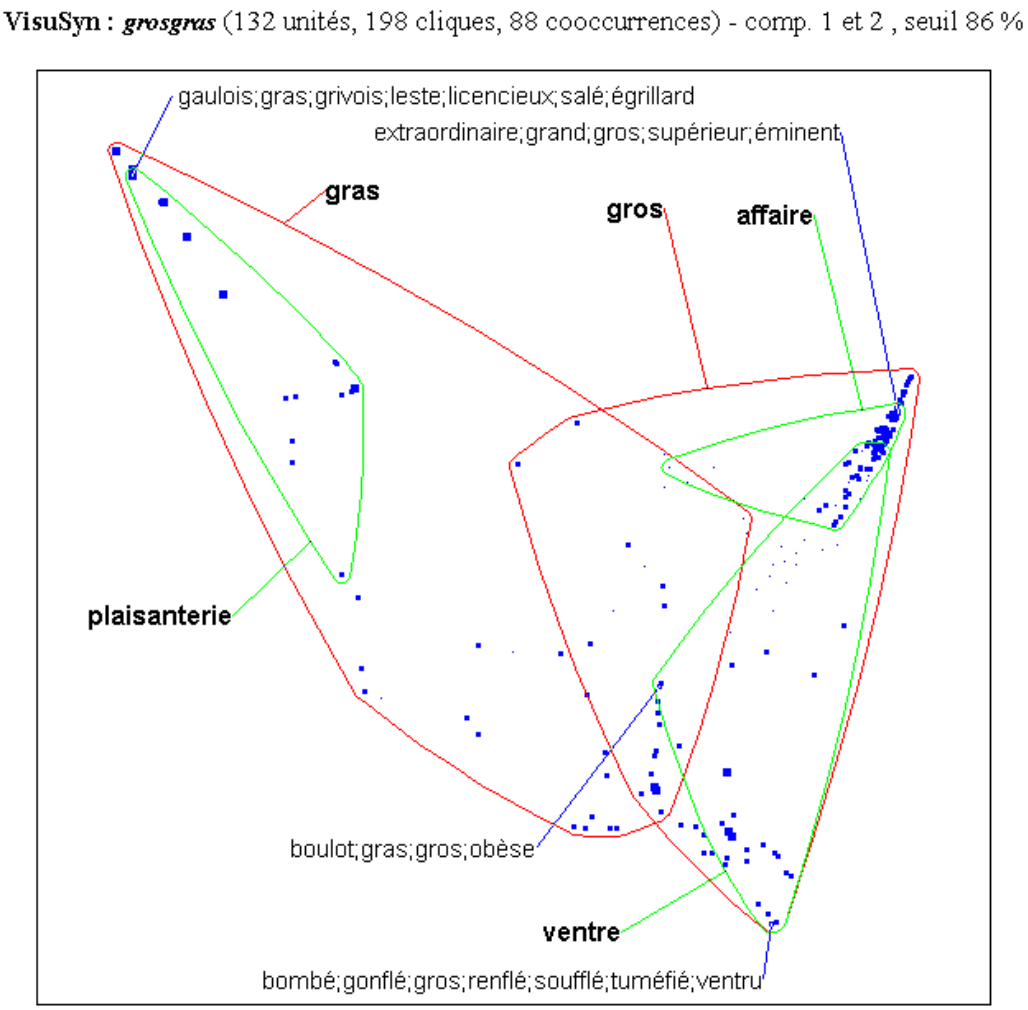


Figure 5 : Représentation de la fonction de compatibilité entre les 198 cliques de gras et gros et trois régissants nominaux : affaire, plaisanterie et ventre[6]

Sur la figure 5 un seuil de 86% de couverture des cliques a été retenu. Supposons que nous souhaitions limiter l'espace de couverture de chacun des trois substantifs aux cliques

[6] Afin de mettre en évidence la zone où gros a une valeur tridimensionnelle ou abstraite — voir la clique {*extraordinaire ; grand ; gros ; supérieur ; éminent*) — un autre plan de visualisation a été adopté pour cette figure.

composées d'adjectifs présentant tous au moins une co-occurrence avec ce substantif. Le taux de couverture est alors de 100%. Pour *plaisanterie* il concerne une clique : {*gras ; grossier ; obscène ; salé*}, pour ventre c'est la clique {*arrondi ; gros ;rebondi ; renflé ; rond*}. Pour *affaire* aucune clique … ne fait l'affaire, le meilleur taux de couverture représenté étant 96% pour {*grave ; gros ; important ; notable ; sérieux*}. Cela signifie qu'il n'y a aucune clique dont tous les constituants aient été rencontrés avec *affaire*, et que l'un des cinq adjectifs de la meilleure clique est absent de la liste de co-occurrences. Cependant le taux n'est pas de 4/5 = 80%, mais de 96% en raison de la méthode de calcul [**A EXPLIQUER SUCCINCTEMENT**].

Cette méthode n'est pas complètement satisfaisante dans la mesure où elle impose la sélection plus ou moins arbitraire d'un seuil et que ce seuil ne peut pas varier pour des relations de co-occurrences avec différents substantifs représentées sur un même graphique. La comparaison dans le tableau 3 entre les cliques retenues comme compatibles avec *plaisanterie* au seuil de 86% (colonne 1), celles qui sont exclues mais ne contiennent que des adjectifs intuitivement[7] compatibles avec ce régissant nominal (colonne 2) et les premières cliques contenant un adjectif incompatible avec plaisanterie (colonne 3) met en évidence le caractère arbitraire de ce seuil.

[7] L'intuition ne devrait jouer aucun rôle dans notre méthode, mais une séquence telle que *plaisanterie grivoise* est absente de la base Frantext catégorisée dont nous nous sommes contentés jusqu'à présent.

cliques retenues au seuil de 86% de la fonction de compatibilité	cliques exclues mais intuitivement entièrement compatibles avec *plaisanterie*	premières cliques contenant un synonyme incompatible avec *plaisanterie*
100 % gras, grossier, obscène, salé 98 % dégoûtant, gras, obscène, répugnant, sale 96 % dégoûtant, gras, grossier, obscène, répugnant 94 % dégoûtant, gras, poisseux, sale 94 % graisseux, gras, sale 92 % gras, obscène, ordurier, répugnant, sale 90 % gras, poisseux, salé 90 % gaulois, gras, grivois, licencieux, obscène, salé 87 % gaulois, gras, grivois, leste, licencieux, salé, égrillard 87 % gras, graveleux, grossier, obscène, ordurier	85 % gras, grossier, obscène, ordurier, répugnant 85 % gras, grossier, obscène, ordurier, répugnant 84 % dégoûtant, gras, grivois, licencieux, obscène, répugnant 82 % gras, graveleux, grossier, malsonnant, ordurier 79 % gros, grossier, lourd, ordinaire 79 % gaulois, gras, graveleux, rabelaisien, égrillard	78 % **adipeux**, graisseux, gras 75 % gras, onctueux, **replet**

Tableau 3 : Cliques retenues et exclues par le seuil de 86% de compatibilité entre les cliques et le régissant nominal *plaisanterie*

Une représentation échelonnée sous forme de "courbes de niveau" de la fonction de compatibilité entre les cliques adjectivales et un régissant nominal a été testée dans François, Victorri & Manguin [à paraître, cf. Figures 2 et 3]. Pour des raisons techniques, elle n'a pas encore pu être appliquée ici en raison de la prise en compte simultanée de plusieurs co-occurents nominaux.

4. Bilan : les particularités respectives de l'espace sémantique de *curieux* et de *gras* / *gros*

Compte tenu de l'incidence sémantique de la position de l'adjectif avant ou après le régissant nominal, l'étude de l'espace sémantique de *curieux* [François, Manguin & Victorri, à paraître] nous a posé un problème spécifique, celui de sa désambiguïsation en position post-nominale. Ici rien de tel : *gros* et *gras* peuvent apparaître dans les deux positions[8] mais sans incidence sémantique notable (contrairement par ex. par exemple à la paire distinctive *un homme grand*

[8] *Gras* se rencontre cependant beaucoup moins fréquemment que *gros* en antéposition (24 occurrences de *grasse matinée* et 7 de *gras pâturages*).

vs. *un grand homme*). En revanche la base textuelle révèle des séquences figées dans lesquelles *gros* ou *gras* n'est pas substituable par un synonyme, ex. *mardi gras*, *foie gras*, *gros do*s /[*faire le ~*] .

L'étude combinée des espaces sémantiques de *gras* et *gros* nous confronte à un autre type de problème. Dans notre base de données textuelles, *gros* et *gras* présentent une co-occurrence au moins avec 23 régissants nominaux :

> *bras, caractère, coeur, dame, doigt, dos, face, femme, figure, fille, garçon, goutte, homme, joue, lèvre, main, monsieur, mot, rire, soulier, trait, voi*x.

Parmi ces substantifs, rares sont ceux dans le contexte desquels *gras* et *gros* sont synonymes : cela vaut pour *caractère* (*écrire en gros caractères / en caractères gras*) et *trait* (*tracer un gros trait / un trait gras*) qui tous deux touchent l'épaisseur d'un graphisme dans un espace bidimensionnel, ce qui neutralise la tridimensionnalité de *gros*. On peut également imaginer qu'un doigt, une main ou une lèvre puissent être qualifiés plus ou moins indifféremment de *gros* ou de *gras*. Dans tous les autres cas *gros* et *gras* ne sont pas synonymes. Cela touche particulièrement les contextes nominaux où *gras* indique la présence d'un film de graisse (ex. *une grosse goutte* vs. u*ne goutte grasse*, *une grosse lettre* vs. *une lettre grasse*, *un gros soulier* vs. *un soulier gras*). Cependant les cliques qui comprennent à la fois *gras* et *gros* dégagent une **acception commune** aux deux adjectifs, par ex. {*abondant, fertile, gras, gros*} où l'on peut soupçonner la sous-jacence d'une relation de métonymie : une herbe **grasse** fournit une **grosse** récolte, un porc **gras** fournit de **gros** jambons, ou {*arrondi, gras, gros, potelé, rebondi, rond*} qui s'applique de préférence à des membres tels que *bras*, *jambes*, *doigts* ou *orteils* dont la grosseur est associée à une présence importante de graisse. La méthode de calcul de la distance entre les cliques constitutives de l'espace sémantique respectif des adjectifs *gras* et *gros* permet donc clairement de dégager l'intersection entre les deux espaces comme un **sous-espace de réelle substituabilité** en contexte nominal.

Bibliographie

Bailly René (1947), *Dictionnaire des synonymes*, Larousse.
Bénac Henri (1956), *Dictionnaire des synonymes*, Hachette.
Bertaud du Chazaud Henri (1989), *Dictionnaire des synonymes*, Robert.
François J., Victorri B., Manguin J.L.(à paraître) Polysémie adjectivale et synonymie : L'éventail des sens de *curieux* étudié à l'aide de méthodes informatiques combinant un dictionnaire électronique de synonymes et des extraits de corpus
Grand Larousse de la Langue Française, Larousse (1971).

Guizot François (1864), *Dictionnaire Universel des synonymes de la Langue Française*, Didier, (7è édition).
Kahlmann A. (1975), *Traitement automatique d'un dictionnaire de synonymes*, Thèse de l'Université de Stockholm.
Lafaye Pierre-Benjamin (1858), *Dictionnaire des synonymes de la Langue Française*, Hachette.
Manguin J.L. (2001), « Construction d'espaces sémantiques associés aux verbes de déplacement d'objets à partir des données des dictionnaires informatisés des synonymes », *Syntaxe et Sémantique*, 2, 287-300.
Manguin J.L.& Victorri B. (1999), « Représentation géométrique d'un paradigme lexical », *TALN 1999*, 363-368.
Ploux S. & Victorri B. (1998), « Construction d'espaces sémantiques à l'aide de dictionnaires de synonymes », *Traitement automatique des langues*, 39-1, 161-182.
Rey, A (éd. 1992). *Dictionnaire historique de la langue françai*se. Editions Le Robert.
Le Grand Robert. Dictionnaire de la Langue Française, Editions Le Robert (1985).