



**HAL**  
open science

## La réduction de la polysémie adjectivale en cotexte nominal : une méthode de sémantique calculatoire

Jacques François, Jean-Luc Manguin, Bernard Victorri

### ► To cite this version:

Jacques François, Jean-Luc Manguin, Bernard Victorri. La réduction de la polysémie adjectivale en cotexte nominal : une méthode de sémantique calculatoire. 2003. hal-01870504

**HAL Id: hal-01870504**

**<https://hal.science/hal-01870504>**

Submitted on 7 Sep 2018

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

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



**CRISCO**

*Centre de Recherches Interlangues  
sur la Signification en Contexte*

**UMR 6170**



**Cahier du CRISCO**

**n°14**

**septembre 2003**

Jacques FRANÇOIS  
Jean-Luc MANGUIN  
Bernard VICTORRI

**La réduction de  
la polysémie adjectivale  
en cotexte nominal :  
une méthode de  
sémantique calculatoire**

**CRISCO**

Université de Caen (Bât. Sciences Porte SA S13), 14032 CAEN CEDEX  
Tél. : 02 31 56 56 27 — Fax : 02 31 56 54 27 — Site web : [www.crisco.unicaen.fr](http://www.crisco.unicaen.fr)  
Courriel direction : [jacques.francois@crisco.unicaen.fr](mailto:jacques.francois@crisco.unicaen.fr)  
Courriel secrétariat : [greneche@crisco.unicaen.fr](mailto:greneche@crisco.unicaen.fr)

**Jacques FRANÇOIS**  
Université de Caen,  
CRISCO (UMR 6170)  
jacques.francois@crisco.unicaen.fr

**Jean-Luc MANGUIN**  
CRISCO (UMR 6170 Caen)  
manguin@crisco.unicaen.fr

**Bernard VICTORRI**  
LATTICe (UMR 8094,  
ENS Montrouge)  
bernard.victorri@ens.fr

## La réduction de la polysémie adjectivale en cotexte nominal : une méthode de sémantique calculatoire

|   |           |
|---|-----------|
| <b>Avertissement</b>  | <b>1</b>  |
| <b>Section 1. Polysémie lexicale et construction du sens en cotexte</b>   | <b>2</b>  |
| <b>Section 2. L'espace sémantique de l'item <i>sec</i> en cotexte nominal</b>   | <b>6</b>  |
| 2.1. Un espace sémantique présentant une « ressemblance de famille »  | 6         |
| 2.2. Construction de l'espace sémantique de l'item <i>sec</i>   | 6         |
| 2.3. Calcul du sens de <i>sec</i> dans une phrase donnée  | 9         |
| <b>Section 3. L'espace sémantique de <i>curieux</i> en cotexte nominal</b>  | <b>13</b> |
| 3.1. Un <i>curieux</i> adjectif pour des linguistes <i>curieux</i>  | 13        |
| 3.2. La polysémie de <i>curieux</i> mesurée à l'aune de la liste de ses synonymes<br>fournie par le <i>Dictionnaire Electronique des Synonymes</i>      | 14        |
| 3.3. Visualisation de la polysémie de <i>curieux</i> à partir de la répartition des cliques<br>de synonymes dans un espace sémantique multidimensionnel | 15        |
| 3.4. Analyse de l'influence du régissant nominal sur le sens de <i>curieux</i> à l'aide de<br>la base FRANTEXT  | 16        |
| 3.5. Prise en compte d'un contexte plus large pour la désambiguïsation des<br>emplois postposés de <i>curieux</i>                                       | 19        |
| 3.6. Bilan  | 23        |
| <b>Section 4. Deux adjectifs occasionnellement synonymes : <i>gros</i> et <i>gras</i> en cotexte<br/>        nominal</b>                                | <b>24</b> |
| 4.1. <i>Gros</i> et <i>gras</i> sont occasionnellement synonymes  | 24        |
| 4.2. La représentation de l'espace sémantique multidimensionnel de <i>gros</i> et de<br><i>gras</i>   | 28        |
| 4.3. L'utilisation de la base Frantext catégorisée et l'ajout du contexte dans la<br>représentation précédente  | 32        |
| 4.4. Bilan  | 36        |
| <b>Section 5. Conclusion</b>  | <b>37</b> |
| <b>Bibliographie</b>  | <b>38</b> |

**Jacques FRANÇOIS**  
Université de Caen,  
CRISCO (UMR 6170)  
jacques.francois@crisco.unicaen.fr

**Jean-Luc MANGUIN**  
CRISCO (UMR 6170 Caen)  
manguin@crisco.unicaen.fr

**Bernard VICTORRI**  
LATTICe (UMR 8094,  
ENS Montrouge)  
bernard.victorri@ens.fr

## La réduction de la polysémie adjectivale en cotexte nominal : une méthode de sémantique calculatoire

### Avertissement

Le présent Cahier 14 du CRISCO se propose de rassembler les résultats de la composante linguistique du projet de recherche « *L'espace sémantique adjectival en contexte discursif : modélisation informatique, analyse de corpus et expérimentation psycholinguistique* » mené par les laboratoires CRISCO (UMR 6170), LATTICe (UMR 8094) et LaCo (UMR 6096) en 2001 et 2002 dans le cadre du programme COGNITIQUE (Thème : *Langage & cognition*) du Ministère de la Recherche. Une partie des résultats de la composante psycholinguistique — point (iii) ci-dessous — est présentée dans François, Manguin, Cordier & Ros (à paraître).

Le triple objectif du projet était

- (i) de calculer la déformation de l'espace sémantique associé à un adjectif par le *Dictionnaire Electronique des Synonymes* du CRISCO (développé par S. Ploux [ISC Lyon], B. Victorri [LATTICe Paris] et J.L. Manguin [CRISCO Caen]) dans un contexte de rection nominale ;
- (ii) de classer un ensemble d'adjectifs français polysémiques et de haute fréquence selon le degré et le type de déformation de l'espace sémantique qui leur est associé en fonction du contexte de rection nominale et
- (iii) d'évaluer expérimentalement sur des sujets francophones natifs d'âge et de formation différente l'accessibilité et le poids relatif de relations synonymiques adjectivales hors contexte et dans différents contextes de rection nominale.

L'étude a porté sur dix adjectifs, dont trois, *curieux*, *gras* et *gros*, ont donné lieu à deux études détaillées en cours de publication<sup>1</sup> qui sont respectivement à la base des sections 3 et 4 du Cahier. La section consacrée à *sec* a donné lieu à des communications, notamment à l'Institut de Psychologie de Lyon 2 (J. François, février 2001) et au colloque *Contrastive Analysis and Linguistic Theory* de Gand (B. Victorri, septembre 2001). L'objectif du Cahier est de permettre une comparaison immédiate des résultats obtenus sur ces quatre adjectifs sur le site du CRISCO.

---

<sup>1</sup> Manguin J.L., François J., Victorri B. (2003): Polysémie adjectivale et rection nominale: quand *gros* et *gras* sont synonymes, in François J. (ed.), *L'adjectif en français et à travers les langues*, Presses Universitaires de Caen.

et

François J., Victorri B., Manguin J.L. (2002): Polysémie adjectivale et synonymie: l'éventail des sens de *curieux*, in Soutet O. (ed.) *La polysémie*, Paris, Presses de l'Université de la Sorbonne.

## Section 1. Polysémie lexicale et construction du sens en cotexte

La polysémie est un phénomène qui se rencontre partout dans les langues. Les mots lexicaux (noms, verbes, adjectifs, etc.) ont rarement une signification unique. Plus communément, un mot peut s'associer à un vaste ensemble de sens différents interconnectés et les items les plus fréquents sont généralement les plus polysémiques. La même remarque s'applique aux marqueurs grammaticaux (marques de temps, prépositions, verbes modaux, etc.). De ce fait, la polysémie ne peut pas être considérée comme un défaut marginal dans les langues. Au contraire, c'est une qualité importante dont l'efficacité est à expliquer.

Dans la plupart des approches classiques, le processus de compréhension d'une phrase est réduit à un simple calcul ascendant (« *bottom-up* ») gouverné par le principe de compositionnalité. Au niveau de base, le sens des syntagmes élémentaires est calculé à partir du sens des unités et le processus est réitéré au long de l'arbre syntaxique, le sens de chaque syntagme à un niveau supérieur étant calculé à partir du sens des syntagmes qui le composent au niveau immédiatement inférieur. Mais un tel processus n'est pas plausible cognitivement : la polysémie serait une charge importante dans ce cas, puisque tous les sens d'une unité polysémique devraient être évalués avant de repérer celui qui convient dans le cotexte et cela conduirait à une explosion combinatoire.

En prenant pleinement en compte le phénomène de la polysémie, nous suggérons une conception différente du processus de construction du sens appelé la *compositionnalité gestaltiste* (Victorri & Fuchs 1996). Le sens d'une unité polysémique est indéterminé au début du processus de construction du sens. Il dépend du sens des autres unités linguistiques dans la phrase (le « cotexte »). Le sens global de la phrase est de ce fait le résultat d'un processus dynamique dans lequel des unités linguistiques interagissent, si bien qu'il se dégage un sens plus ou moins précis pour chaque unité aussi bien qu'un sens global pour la phrase entière.

Le modèle appliqué dans ce projet est destiné à capturer la nature dynamique de ce processus par le biais de la théorie mathématique des systèmes dynamiques. Ce modèle est implémenté sur un ordinateur et articulé sur deux ressources linguistiques électroniques, un vaste dictionnaire des synonymes du français et un corpus important de textes.

En premier lieu, on peut représenter l'espace sémantique — c'est-à-dire la polysémie ou dans la terminologie de la sémantique lexicale européenne, le « champ sémasiologique » — d'une unité lexicale sous la forme d'un espace sémantique présentant autant de dimensions que de synonymes qui lui sont associés (il s'agit en général de synonymes partiels, c'est-à-dire relatifs à des contextes discursifs particuliers). C'est la démarche adoptée par S. Ploux et B. Victorri (1998) dans la constitution du *Dictionnaire Electronique des Synonymes* (désormais abrégé D.E.S.)<sup>2</sup>. Un logiciel, VISUSYN, fournit un jeu de visualisations de cet espace sémantique multidimensionnel par plans sur lesquels sont projetés sous forme de points des cliques de synonymes, c'est-à-dire des ensembles d'unités en relation de renvoi synonymique mutuel dans une base constituée des renvois synonymiques de sept dictionnaires<sup>3</sup>. 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 se révèle a posteriori apparentée au travail effectué il y a un quart de siècle par André Kahlmann<sup>4</sup>. La base est constituée de

---

<sup>2</sup> Pour une application au lexique verbal, cf. Manguin (1999, 2001).

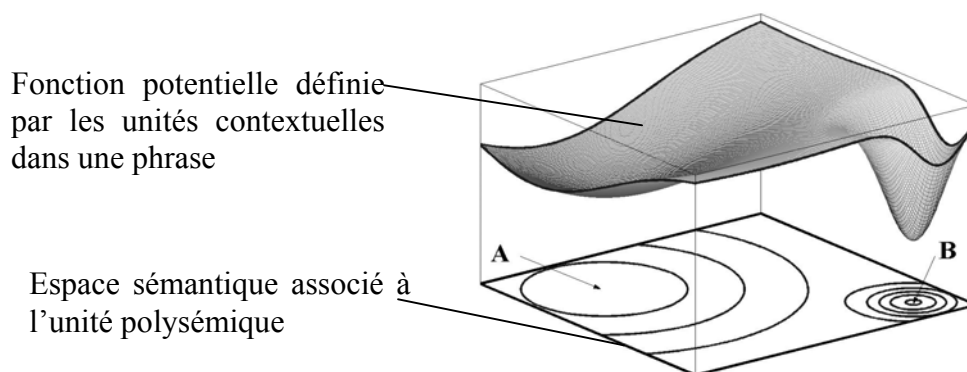
<sup>3</sup> Le D.E.S. est en démonstration sur le site <http://www.crisco.unicaen.fr/dicosyn.html> du Centre de Recherches Inter-Langues sur la Signification en Contexte (Université de Caen et CNRS).

<sup>4</sup> cf. Kahlmann A. (1975), où était proposée la constitution d'un dictionnaire électronique par cumul des entrées et des relations synonymiques de plusieurs dictionnaires de synonymes existant en version papier.

deux dictionnaires de langues extensifs, le *Grand Larousse de la Langue Française* (1971) et le grand *Dictionnaire Robert* (1992), de quatre dictionnaires de synonymes explicatifs, deux du 19<sup>e</sup> siècle, celui de Guizot (1864) et celui de Lafaye (1878) et deux du 20<sup>e</sup>, le dictionnaire de Bénac (1956) et celui de Bailly (1947) 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 (seul Du Chazaud ne classe pas les synonymes), c'est-à-dire que les relations de synonymie sont décontextualisées.

Dans ces représentations, il n'est pas tenu compte du contexte discursif, seul le point de vue paradigmatique est donc en cause. Cependant on sait que l'espace sémantique d'un lexème résulte de l'union de sous-espaces sémantiques correspondant à l'emploi de ce lexème dans les différents types de contexte discursif où il peut apparaître. Dans le cas des adjectifs, un lexème adjectival Adj<sub>1</sub> présentant comme synonymes partiels {Adj<sub>2</sub>, Adj<sub>3</sub>, Adj<sub>4</sub>, etc.} et susceptible d'apparaître comme épithète ou attribut des régissants nominaux {N<sub>1</sub>, N<sub>2</sub>, N<sub>3</sub>, etc.} peut par exemple être remplaçable par Adj<sub>2</sub> comme épithète postposé de N<sub>1</sub>, par Adj<sub>3</sub> comme épithète antéposé de N<sub>2</sub> et par Adj<sub>4</sub> comme attribut de N<sub>1</sub><sup>5</sup>. Il est donc souhaitable de se donner les moyens de calculer l'effet du contexte discursif sur l'espace sémantique associé à un lexème.

Le calcul du sens d'une unité polysémique dans une phrase donnée est modélisé comme un processus dynamique. Les unités contextuelles pertinentes dans la phrase définissent une « fonction potentielle » sur l'espace sémantique (cf. Figure 1.1). Les valeurs potentielles en deçà d'un seuil donné déterminent une région de l'espace sémantique qui représente le sens de l'unité dans la phrase en question. Cette région est constituée des « bassins d'attraction » des dynamiques dérivant de la fonction potentielle<sup>6</sup>.



**Figure 1.1:** Représentation d'une fonction potentielle sur un espace sémantique bidimensionnel

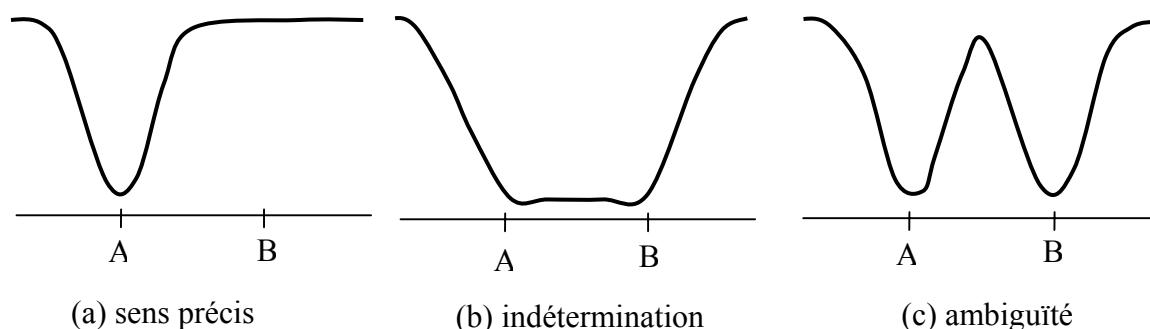
Comme on peut le voir sur la Figure 1.2, il est possible de représenter dans le même cadre l'éventail complet des types d'interprétation (sens précis, ambiguïtés et indéterminations)<sup>7</sup>. Un sens précis correspond à une région étroite (a), une indétermination à une région plus vaste

<sup>5</sup> Nous faisons provisoirement abstraction des emplois phraséologiques.

<sup>6</sup> Pour tous les détails techniques, voir Victorri & Fuchs 1996, spécialement au chapitre 3 : *Un modèle morphodynamique* (pp. 63-85). La fonction de compatibilité entre un nom régissant et un adjectif régi (cf. §3.4) est un cas particulier de cette fonction potentielle.

<sup>7</sup> Pour des représentations apparentées, voir Sadock (1986).

s'étendant à deux sens ou plus (b) et une ambiguïté à une région constituée de deux parties déconnectées ou plus (c). Ainsi le modèle implémente un mécanisme général pour la construction du sens d'items polysémiques qui est plus plausible d'un point de vue cognitif.



**Figure 1.2:** Différents schèmes correspondant à différents types d'interprétation dans un espace sémantique unidimensionnel

La relation nominale des adjectifs (épithètes et attributs) constitue l'une des relations de contextualisation discursive les plus élémentaires, c'est pourquoi nous nous proposons d'aborder le problème par cette voie<sup>8</sup>. Nous limitant provisoirement aux paires [Nom régissant, Adjectif épithète / apposé] nous avons étudié comment la présence de ce nom contraint le sens de l'adjectif, ce qui a fait jusqu'à présent l'objet d'une évaluation grossière pour les adjectifs *beau*, *difficile*, *doux*, *froid*, *propre*<sup>9</sup> et *triste* et minutieuse pour les quatre adjectifs mentionnés plus haut.

Pour ce faire, il nous faut recourir à une base de données textuelle à la fois vaste et catégorisée. La base FRANTEXT catégorisée, même si elle présente l'inconvénient de ne pas prendre en compte le discours journalistique et de ne pas couvrir les années postérieures à 1970<sup>10</sup>, satisfait cette exigence. Ce corpus nous permet de rechercher des éléments en fonction de leur catégorie grammaticale (ce qui est impossible dans la base Frantext non catégorisée). L'objectif est d'obtenir une liste de paires {Nom, Adjectif}, où chaque Adjectif est un synonyme de l'item étudié, et où chaque Nom co-occure au moins une fois avec celui-ci.

Pour obtenir cette liste, nous procédons en deux étapes :

(i) Tout d'abord nous cherchons les noms en cooccurrence avec l'adjectif, par exemple *curieux*, par deux requêtes Frantext données ci-dessous :

<sup>8</sup> L'analyse de l'écart entre les cliques de synonymes constitutives de l'espace sémantique d'un item polysémique est purement descriptive et ne dit rien des processus de diversification du sens. Depuis Darmesteter (1887/1979, pp. 67-75), on sait qu'historiquement la diversification du sens s'effectue de manière radiale (processus de « rayonnement » selon Darmesteter) ou linéaire (processus d'« enchaînement » selon Darmesteter), entrant fréquemment en combinaison. On peut se reporter à Nyckees (1998, pp. 89-156) et Niklas-Salminen (1996, pp.159-165) sur les modalités de l'évolution du sens. La réalité cognitive de la polysémie, en particulier du type « ressemblance de famille » selon L. Wittgenstein, est sujette à controverse et l'exemple célèbre de la polysémie de l'item *game* en anglais (Wittgenstein 1961, pp. 147-8) a donné lieu à des développements intéressants à comparer de Nyckees (1998, pp. 320-7) et Touratier (2000, pp. 110).

<sup>9</sup> L'espace sémantique de l'item adjectival et nominal *propre* a été analysé dans François & Manguin (à paraître) mais l'étude de la réduction de cet espace en co-texte nominal n'est pas encore achevée.

<sup>10</sup> A l'époque où la recherche des co-occurrences a été menée. Dans les derniers mois, des textes plus récents ont été catégorisés par l'ATILF.

$\&e(g=S)$        $\&?(\&e(g=Adv))$        $\&e(c=(\&mcurieux) g=A)$       (A)  
 ↑                            ↑                            ↑  
 un substantif      et un adverbe éventuel      et un adjectif, forme fléchie de *curieux*

$\&?(\&e(g=Adv))$        $\&e(c=(\&mcurieux) g=A)$        $\&e(g=S)$       (B)  
 ↑                            ↑                            ↑  
 un adverbe éventuel      et un adjectif, forme fléchie de *curieux*      et un substantif

L'analyse des réponses de Frantext à l'aide d'un programme nous donne une liste de noms (que nous appellerons *liste\_noms*). Une requête sur notre dictionnaire des synonymes nous avait déjà donné une liste d'adjectifs synonymes (en l'occurrence de *curieux*), que nous appellerons *liste\_adjectifs*.

(ii) 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)  
 ↑                            ↑                            ↑  
 un substantif de la      et un adverbe éventuel      et un adjectif de la liste  
 liste "liste\_noms"      "liste\_adjectifs"

$\&?(\&e(g=Adv)) \&e(c=\&lliste\_adjectifs g=A) \&e(g=\&lliste\_noms)$       (B)  
 ↑                            ↑                            ↑  
 un adverbe éventuel      et un adjectif de la      et un substantif de la liste "liste\_noms"  
    liste "liste\_adjectifs"

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.

| Adjectif     | Nb de cooccurrences |
|--------------|---------------------|
| beau         | 58057               |
| gros         | 49595               |
| grand        | 46566               |
| doux         | 31837               |
| propre       | 23021               |
| sec          | 16562               |
| triste       | 13981               |
| difficile    | 13497               |
| curieux      | 11469               |
| froid        | 10273               |
| <b>Total</b> | <b>274858</b>       |

**Tableau 1.1** : Nombre de cooccurrences relevées dans le corpus Frantext



## Section 2. L'espace sémantique de l'item *sec* en cotexte nominal

### 2.1. Un espace sémantique présentant une « ressemblance de famille »

L'adjectif *sec* est notoirement hautement polysémique. Le *TLF* (Quemada 1992) en fournit plus de trente acceptions, dont les six principales sont :

- |  |   |
|--|---|
| (1) dénué d'eau,                           | ex. <i>La route était sèche</i>             |
| (2) mince, décharné,                       | ex. <i>Un vieil homme sec et ridé</i>       |
| (3) stérile, improductif,                  | ex. <i>Rester sec à un examen</i>           |
| (4) dénué d'émotion, de douceur, égoïste,  | ex. <i>un homme au cœur sec.</i>            |
| (5) bref, abrupt, manquant de délicatesse, | ex. <i>donner un coup sec</i>               |
| (6) isolé, seul,                           | ex. <i>avoir un atout sec dans son jeu.</i> |

Bien que ces sens soient très différents les uns des autres, il est aisé de voir qu'ils peuvent être reliés les uns aux autres. Les sens 1,2 et 3 sont liés quand *sec* s'applique à des plantes ou de la végétation. Le lien entre les sens 3 et 4 est également clair: une personne égoïste est quelqu'un qui ne donne pas beaucoup. Le sens 5 qui s'applique à des événements, entretient une relation avec le sens 4 en caractérisant un comportement mal dégrossi. Finalement le sens 6 peut être relié au 5 autant qu'aux sens 2 et 3: un atout *sec* est une carte manquant de « compagnie » pour être pleinement profitable.

Ainsi l'ensemble complet des sens se laisse décrire comme une « ressemblance de famille » (Wittgenstein 1961, Lakoff 1987) dans le cadre théorique que Kleiber (1990) appelle « la théorie étendue du prototype ». La théorie présuppose que les différents sens d'un simple item sont pourvus d'une relation de distance : deux sens sont proches s'ils partagent une relation forte et ils sont plus éloignés s'ils partagent moins de traits sémantiques.

Il est important de noter que chacun des six sens prototypiques de *sec* est en fait une collection de « nuances » différentes. Par exemple, pour le sens 4 nous pouvons distinguer différentes qualités psychologiques de l'égoïsme à l'insensibilité et à la rudesse. A cela s'ajoute le fait que plus d'un sens peut être impliqué dans une phrase donnée. Ainsi le syntagme *un ton sec* caractérise à la fois les aspects physiques (sens 5) et psychologique (sens 4) d'un ton. Un autre exemple est fourni par *une terre sèche* qui dénote un terrain à la fois desséché (sens 1) et stérile (sens 2). Ces cas que nous qualifierons d'indéterminations, sont très fréquents. Bien qu'ils soient souvent négligés dans les analyses sémantiques, ils sont très intéressants d'un point de vue cognitif, parce qu'ils peuvent expliquer pourquoi la polysémie est efficace : très souvent le locuteur veut exprimer une pensée complexe qui combine deux ou plusieurs idées corrélées et en recourant à des items polysémiques, ce besoin peut être satisfait avec une efficacité accrue. Ces cas doivent être distingués d'ambiguïtés réelles dans lesquelles deux sens sont aussi impliqués, mais où seul l'un des deux est pertinent. Par exemple, dans le syntagme *un homme sec*, l'adjectif peut avoir un sens physique (un homme mince) ou psychologique (un homme insensible), mais seul l'un des deux sens est visé dans un contexte donné.

### 2.2. Construction de l'espace sémantique de l'item *sec*

Différentes méthodes sont disponibles pour construire l'espace sémantique associé à des items linguistique. Pour les marques grammaticales, une analyse linguistique est nécessaire afin de repérer (1) le sens élémentaire (ou "forme schématique") commun à tous les sens de

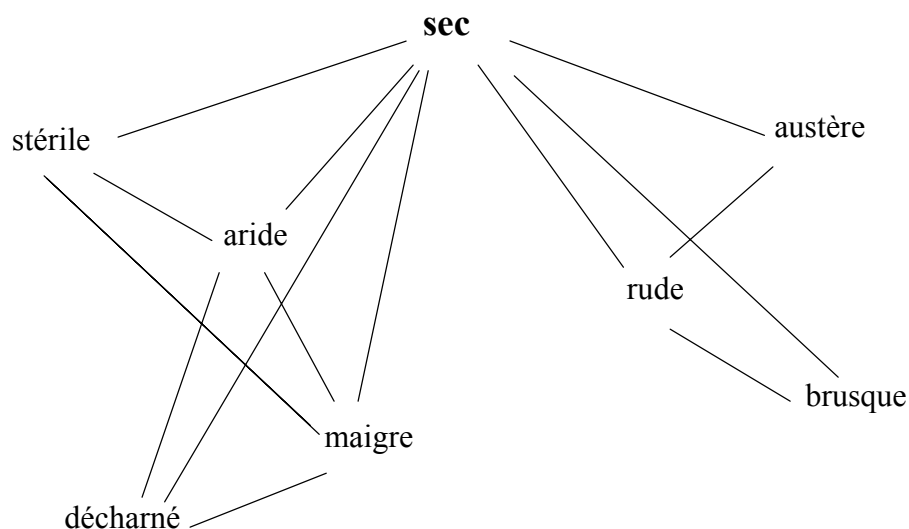
l'item et (2) l'ensemble des paramètres rendant compte des variations entre ces sens. Dans Victorri & Fuchs (1996), on montre comment appliquer cette méthode pour construire un espace sémantique associé à l'adverbe français aspectuel et modal *encore*. Dans cette section nous nous concentrerons sur la méthode exposée en Section 1 et nous l'illustrerons à l'aide de l'exemple de l'adjectif *sec*.

Le *D.E.S.* fournit 56 synonymes de *sec*. En voici la liste :

*aigre, âpre, aride, austère, autoritaire, blessant, bourru, bref, brusque, brutal, cassant, concis, cru, décharné, déplaisant, désagréable, désertique, désobligeant, desséché, dur, efflanqué, égoïste, émacié, endurci, étique, étriqué, fauché, ferme, froid, glacé, glacial, improproductif, indifférent, ingrat, insensible, maigre, maigrelet, pauvre, racorni, raide, rébarbatif, rebutant, revêche, rigide, rogue, rude, sec, séché, seul, sévère, simple, squelettique, stérile, tranchant, vert, vide.*

Il importe de noter qu'ils ne sont que partiellement synonymes de *sec*, puisqu'ils ne peuvent pas remplacer *sec* dans tout contexte : chaque synonyme ne partage avec *sec* qu'une part de son espace sémantique.

Le dictionnaire fournit plus d'informations que cette simple liste. Nous pouvons savoir pour toute paire d'items dans la liste s'ils sont synonymes l'un de l'autre ou pas. En d'autres termes, le graphe complet des relations de synonymie est disponible. Ce graphe est naturellement très complexe et ne se laisse pas représenter aisément. La Figure 2.1 montre une petite partie du graphe à laquelle nous allons nous limiter pour expliquer les principes de base de notre méthode.



**Figure 2.1** : Une partie du graphe de synonymie de *sec*

Comme on peut le voir sur la figure 4.3, un simple synonyme ne suffit généralement pas pour définir un sens précis de *sec*. Par exemple, *maigre*, qui est synonyme à la fois de *décharné* et de *stérile*, empiète à la fois sur les sens 2 (*mince*) et 3 (*improductif*). Similairement, *rude* est compatible avec le sens 4 (*manquant de douceur*) et le sens 5 (*manquant de délicatesse*). Mais

si nous considérons ce qui en théorie des graphes<sup>11</sup> est appelé des cliques, c'est-à-dire les sous-ensembles composés des items qui sont mutuellement synonymes, nous constatons que celles-ci définissent des sens plus précis. De fait, le graphe présenté en Figure 2.1 contient les quatre cliques suivantes, chacune correspondant à un sens de *sec* :

- *aride, décharné, maigre, sec* (sens 2)
- *aride, maigre, sec, stérile* (sens 3)
- *austère, rude, sec* (sens 4)
- *Brusque, rude, sec* (sens 5)

En fait, si nous calculons toutes les cliques du graphe complet de synonymie de *sec* (comprenant les 56 synonymes), nous trouvons 75 cliques. Voici un échantillon de la liste des cliques).

- *aride, desséché, décharné, maigre, sec.*
- *aride, désertique, ingrat, sec, stérile.*
- *aride, desséché, maigre, sec, stérile*
- *aride, improductif, pauvre, sec, stérile, vide.*
- *austère, rude, sec, simple, sévère*
- *autoritaire, brusque, cassant, sec, tranchant.*
- *bourru, brusque, brutal, cru, rude, sec.*
- *bref, brusque, brutal, sec.*
- *bref, concis, sec.*
- *dur, froid, indifférent, insensible, sec.*
- *fauché, pauvre, sec.*
- *indifférent, insensible, sec, égoïste.*
- *sec, seul, simple.*
- ...

Il ressort de cet échantillon que chaque clique définit un sens très précis de *sec*, une nuance de sens. Il est alors naturel d'associer chaque clique à un point dans l'espace sémantique. Le pas suivant de notre méthode consiste à définir une distance entre cliques, d'une manière telle que des cliques définissant des sens proches soient représentées par des points proches dans l'espace sémantique. Une fois que l'ensemble des cliques est doté de cette métrique, nous pouvons obtenir des représentations à faible nombre de dimensions. Nous ne donnerons pas ici les détails des algorithmes<sup>12</sup>. Le résultat final pour l'adjectif *sec* est représenté sur la Figure 2.2.

Comme on peut s'en rendre compte sur cette figure, les différents sens de *sec* sont clairement identifiables dans l'espace sémantique. Le sens 1 (dénué d'eau) est localisé au centre de l'espace (représenté par la clique {*desséché, sec, séché*}). En se déplaçant de ce point vers le coin en haut à gauche nous progressons vers le sens 3 (improductif) puis vers le sens 2 (mince). A droite sont localisés les sens psychologiques (sens 4) : en bas on trouve les cliques définissant l'égoïsme et l'insensibilité, et en se déplaçant vers le haut, on accède à la sévérité et à la rudesse. A proximité de cette dernière nuance, nous trouvons en haut à droite

<sup>11</sup> Pour une introduction à la théorie des graphes, voir par exemple Berge (1970). L'algorithme utilisé pour calculer les cliques se trouve dans Reingold *et al.* (1977). Une approche similaire recourant à un graphe de synonymie se rencontre dans Warnesson (1992). Les cliques sont complètement différentes des "synsets" de Wordnet (Fellbaum 1998) : dans Wordnet, les mots sont désambiguïsés avant le calcul des "synsets".

<sup>12</sup> Tous les détails sont fournis dans Ploux & Victorri (1998). Pour des illustrations de l'application de la méthode à d'autres items polysémiques, voir Manguin & Victorri 1999, Manguin 2001, Victorri 2002, François *et al.* 2002 et Manguin *et al.* 2003.

le sens 5 (qualifiant des événements). Finalement le sens 6 (isolé) est localisé en haut au centre.

On peut vérifier que la représentation satisfait nos attentes. Virtuellement toutes les relations sémantiques entre les différents sens de *sec* sont prises en compte en termes de distance entre cliques et la plupart des nuances sont localisées à des places appropriées.

### 2.3. Calcul du sens de *sec* dans une phrase donnée

Pour calculer automatiquement le sens de *sec* dans différentes phrases, nous utilisons un vaste corpus électronique qui nous permet de découvrir comment les éléments contextuels dans la phrase (et en premier lieu le nom régissant cet adjectif) restreignent son sens. Ce travail est encore en cours et nous allons présenter ici nos résultats préliminaires.

A partir du corpus de textes français FRANTEXT<sup>13</sup>, nous récupérons les occurrences de *sec* et nous établissons la liste des principaux noms régissant *sec* dans le corpus. Voici un échantillon de la liste :

*air, arbre, bois, boue, bras, chambre, chemin, cheveu, chose, claquement, cœur, coin, corps, cou, coup, doigt, éclat, esprit, façon, femme, feuille, figure, fleur, foin, fromage, geste, gorge, herbe, homme, jambe, jardin, lèvres, ...*

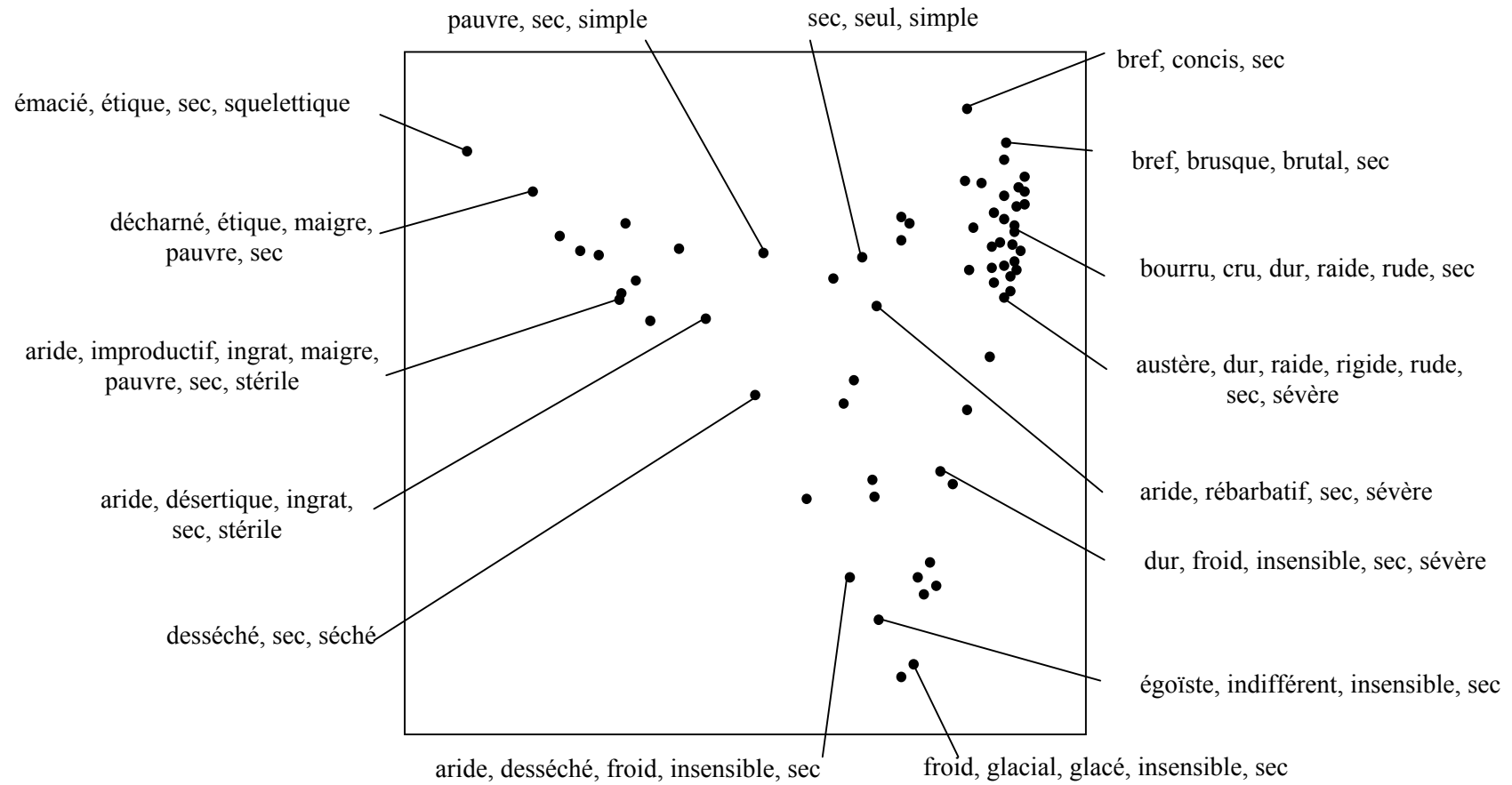
Ensuite nous avons recherché dans le corpus des occurrences de tous les syntagmes contenant un nom de cette liste et un adjectif de la liste des synonymes de *sec*. A partir de toutes ces données, nous avons calculé la « fonction de compatibilité » d'un nom avec une clique. Cette fonction de compatibilité a une valeur élevée s'il y a de nombreuses co-occurrences du nom avec chaque élément de la clique.

Pour illustrer le principe du calcul, prenons un exemple restreint, en l'occurrence le nom terre et le sous-ensemble limité des synonymes introduits dans la Figure 4.3 (*aride, austère, brusque, décharné, maigre, rude, sec, stérile*). Le nombre de co-occurrences trouvées dans le corpus (classé par ordre décroissant) est le suivant :

|                        |   |
|------------------------|---|
| <i>terre aride</i>     | 8 |
| <i>terre stérile</i>   | 6 |
| <i>terre sèche</i>     | 3 |
| <i>terre maigre</i>    | 2 |
| <i>terre austère</i>   | 1 |
| <i>terre rude</i>      | 1 |
| <i>terre brusque</i>   | 0 |
| <i>terre décharnée</i> | 0 |

---

<sup>13</sup> Frantext est accessible sur le site internet [http://www.atilf.fr/ie/index\\_frantext.htm](http://www.atilf.fr/ie/index_frantext.htm). Pour des revues récentes de l'utilisation de corpus en linguistique calculatoire, on peut se reporter par exemple à Habert *et al.* 1997, Pierrel 2000, Bourigault *et al.* 2001. Voir aussi Ide & Véronis 1990 pour une autre approche de la désambiguïsation automatique des sens lexicaux mettant en œuvre un dictionnaire électronique.



**Figure 2.2** : L'espace sémantique associé à *sec*

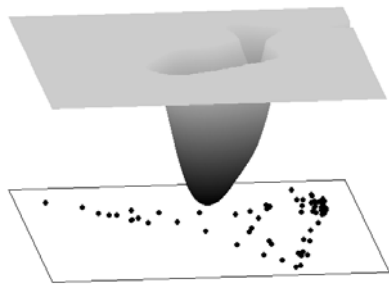
Le nom *terre* est pleinement compatible avec les adjectifs *aride* et *stérile* (représentatifs du sens 3), à un moindre degré avec *maigre*, *rude* et *austère*, et pas du tout avec  *Brusque* et *décharné*. De ce fait, la fonction de compatibilité de *terre* aura sa valeur la plus élevée avec la clique {*aride, maigre, sec, stérile*}, et la plus basse avec la clique { *Brusque, rude, sec*}. Le calcul exact, qui prend en compte la fréquence relative des adjectifs du corpus, fournit les valeurs qui suivent :

|                                     |      |
|-------------------------------------|------|
| <i>sec, aride, maigre, stérile</i>  | 98 % |
| <i>sec, aride, maigre, décharné</i> | 57 % |
| <i>sec, rude, austère</i>           | 42 % |
| <i>sec, rude, brusque</i>           | 28 % |

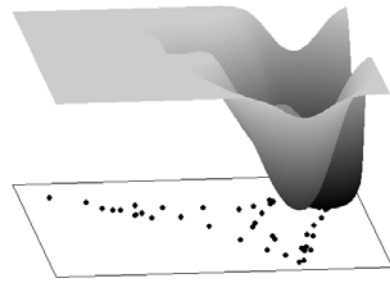
A partir de ces exemple modeste nous pouvons inférer que la fonction de compatibilité pourrait être un bon indicateur de la manière dont le nom régissant influence le sens de *sec*. L'idée sous-jacente est que le nom restreint le sens de l'adjectif à la partie de l'espace sémantique où sa fonction de compatibilité avec les cliques prend une valeur élevée. En d'autres termes, nous pouvons dériver la fonction potentielle associée au nom de sa compatibilité avec les cliques. C'est ce qu'illustre la Figure 2.3 sur laquelle la fonction potentielle a été calculée pour quatre noms : *foin*, *ton*, *homme* et *cœur*.

Les différents types d'interprétation y sont ainsi représentés :

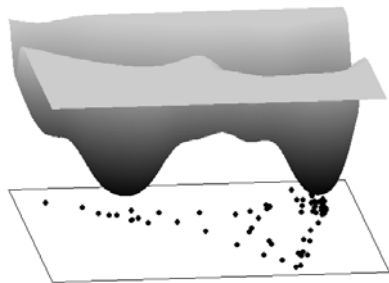
- *foin* (a) donne lieu à une fonction potentielle avec un bassin étroit situé au centre de l'espace sémantique. Cela signifie que *sec* a un sens très précis dans ce contexte, à savoir 1 ('dénué d'eau').
- *ton* (b) est associé à une fonction potentielle présentant un vaste bassin couvrant presque toute la partie droite de l'espace sémantique. C'est un exemple du cas que nous avons désigné comme indétermination : aussi bien le sens 4 ('manquant de douceur') que le sens 5 ('manquant de délicatesse') sont convoqué dans ce contexte.
- *homme* (c) dégage une ambiguïté typique. La fonction potentielle présente deux bassins séparés, l'un couvrant les sens psychologiques de *sec*, l'autre les sens physiques.
- *cœur* (d) est un exemple intéressant. La fonction potentielle présente deux minima, l'un centré sur la clique {*aride, rébarbatif, sec, sévère*} et l'autre sur la clique {*égoïste, indifférent, insensible, sec*}. Cela veut dire que seuls les sens psychologiques de *sec* sont sélectionnés dans ce contexte. Dans la plupart des approches classiques, un tel cas de figure présenterait un problème difficile car aussi bien le sens de *cœur* que celui de *sec* seraient qualifiés de métaphoriques. A titre d'exemple, dans la théorie de Pustejovsky (1995) il serait difficile d'expliquer par un mécanisme de « coercion de type » que les deux unités à la fois passent d'un sens physique à un sens psychologique. Dans notre approche nous ne sommes pas confrontés à ce problème, puisque nous ne dérivons pas les sens supposés métaphoriques d'un sens 'premier' ou 'littéral' quelconque. Chacun des sens est obtenu par le biais du même mécanisme dynamique.



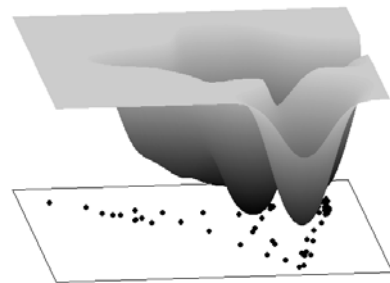
(a) *foin sec*



(b) *ton sec*



(c) *homme sec*



(d) *cœur sec*

**Figure 2.3:** Fonctions potentielles associées à différents noms cooccurrent avec *sec*

### Section 3. L'espace sémantique de *curieux* en cotexte nominal

Comme l'a bien montré Bat-Zeev Schyldkrot (1997), l'adjectif *curieux* occupe une place particulière dans le lexique adjectival du français, en raison de son extrême sensibilité non seulement à la classe conceptuelle des substantifs propres à le régir comme épithète, mais aussi à la place du régissant nominal : [Det N *curieux*] vs. [Det *curieux* N]. Nous nous proposons de montrer que certains contextes de rection nominale avec épithète préposé ou postposé ont le pouvoir de réduire l'espace sémantique associé à cet adjectif par le *D.E.S.*, c'est-à-dire de sélectionner un effet de sens. En revanche, d'autres contextes nominaux n'ont pas le même pouvoir discriminant et dans ce cas la désambiguïsation doit faire appel à la constitution d'une scène mentale qui peut mettre en jeu un contexte très élargi et un savoir extralinguistique étendu. Mais avant d'aller plus loin, il importe de prendre la mesure de la polysémie de *curieux*.

#### 3.1. Un *curieux* adjectif pour des linguistes *curieux*

Le parcours sémantique de *curieux*, tel que le résume le *Dictionnaire Historique de la Langue Française* (A. Rey, éd. 1992) a pour source l'idée de « soin » véhiculée par le substantif latin *cura* dont le dérivé *curiosus* désigne celui qui prend soin, jusqu'à l'indiscrétion, puisque le terme qualifie un espion en latin impérial. La valeur de soin, souci, préoccupation est dominante en ancien français (*curius*, 12<sup>e</sup>), celle d'indiscrétion se développe au 16<sup>e</sup>. Au 17<sup>e</sup> siècle s'ajoute la qualification de celui «qui rassemble, recherche des objets rares ou précieux» (voir les 'cabinets de curiosités') et au 18<sup>e</sup> la valeur péjorative de l'indiscrétion cède la place à celle méliorative de recherche désintéressée. Dans le même temps «pour exprimer la valeur passive de «digne d'intérêt, original, étrange» (1559), l'usage a recours à l'antéposition de l'épithète (*un curieux personnage*) afin d'éviter l'ambiguïté, sauf dans *bête curieuse* (1755)».

Au terme de son enquête historique, Bat-Zeev Schyldkrot (1997:123) conclut à l'émergence d'un couple de valeurs synonymes respectivement de *intéressé* et *intéressant* et l'auteur formule la relation entre la valeur que nous qualifierons d'AGENTIVE (SYN *intéressé*) et la valeur de STIMULUS (SYN *intéressant*) dans des termes apparentés à ceux du *Dictionnaire historique* :

« Le sens 'digne d'intérêt', 'bizarre' s'est développé plus tard mais s'est vite fixé en position antéposée au substantif conformément aux propriétés attribuées à ces adjectifs. L'antéposition permet de lever l'ambiguïté qui existe pour les deux autres sens »

Dans les deux cas, on insiste sur la distinction syntagmatique comme **moyen** de désambiguïsation, mais en fait seul l'ordre Adj - N a cet effet, sélectionnant la valeur 'STIMULUS' (ex. *un curieux personnage*), sans que cette valeur soit exclue dans l'ordre inverse (*un personnage curieux* ≈ *surprenant* vs. ≈ *avide d'apprendre*).



### 3.2. La polysémie de *curieux* mesurée à l'aune de la liste de ses synonymes fournie par le *Dictionnaire Electronique des Synonymes*

La liste des 51 synonymes de *curieux* obtenue en consultant le D.E.S. est constituée, nous l'avons déjà évoqué en Section 1, à partir des renvois synonymiques de sept dictionnaires de langue ou de synonymes<sup>14</sup>.

*amateur, amusant, anxieux, attachant, attentif, avide, beau, bizarre, bizarroïde, chercheur, collectionneur, déconcertant, désireux, dilettante, drolatique, drôle, étonnant, étrange, extraordinaire, extravagant, flâneur, fouille-au-pot, fouilleur, fouinard, fouineur, fureteur, incompréhensible, inconcevable, incroyable, indiscret, inouï, inquisiteur, intéressant, intéressé, inusité, investigateur, original, paradoxal, piquant, pittoresque, plaisant, rare, sensationnel, singulier, soigneux, soucieux, spectateur, surprenant, truculent, unique, voyeur*

En dehors des deux valeurs AGENT [A] et STIMULUS [S] identifiées ci-dessus, une troisième valeur mérite d'être distinguée pour les emplois de *curieux* régissant un complément prépositionnel à valeur prédicative et en général à l'infinitif (ex. *je suis curieux de [connaître] sa réaction / d'en savoir plus*). *Curieux* et ses synonymes dans cette construction dénotent une DISPOSITION [D] à accomplir le type d'action exprimé par la construction infinitive.

L'examen de la relation de synonymie entre ces adjectifs et *curieux* donne lieu à plusieurs observations :

- (I) Seul de tous ces adjectifs à pouvoir prendre plus d'une valeur — et de fait les trois [A,D,S] — *curieux* ouvre un champ sémasiologique qui présente trois groupes de synonymes respectivement caractérisés par l'une des trois valeurs d'AGENT, de DISPOSITION ou de STIMULUS :

#### *curieux*<A>

SYN<sup>15</sup> : *amateur, chercheur, collectionneur, dilettante, flâneur, fouille-au-pot, fouilleur, fouinard, fouineur, fureteur, indiscret, inquisiteur, investigateur, spectateur, voyeur*

#### *curieux*<D>

SYN<sup>16</sup> : *anxieux de INF, attentif à INF, avide de INF, désireux de INF, intéressé à INF, soigneux de INF [vx], soucieux de INF*

#### *curieux*<S>

SYN : *amusant, attachant, beau, bizarre, bizarroïde, déconcertant, drolatique, drôle, étonnant, étrange, extraordinaire, extravagant, incompréhensible, inconcevable, incroyable, inouï, intéressant, inusité, original, paradoxal, piquant, pittoresque, plaisant, rare, sensationnel, singulier, surprenant, truculent, unique*

- (II) L'ensemble des adjectifs qui peuvent être préposés à leur régissant nominal est un sous-ensemble de ceux du type S :

<sup>14</sup> La liste des synonymes de *curieux* fournie par le D.E.S. comporte également *badaud, regardant* et *juge* qui n'entrent en relation de synonymie qu'avec des emplois nominaux de *curieux* et qui sont écartés ici.

<sup>15</sup> *Amateur* et *collectionneur* sont fréquemment suivis d'un complément prépositionnel mais à statut non prédicatif.

<sup>16</sup> Il est à noter que *anxieux* et *soucieux* n'entretiennent de relation de synonymie avec *curieux* que lorsqu'ils introduisent un complément prédicatif à l'infinitif

— synonymes du type [S] préposables :

*amusant, attachant, beau, bizarre, bizarroïde, drôle [de N], étonnant, étrange, extraordinaire, extravagant, incompréhensible, inconcevable, incroyable, intéressant, paradoxal, piquant, pittoresque, plaisant, rare, sensationnel, singulier, surprenant, truculent, unique (?)*

— synonymes du type [S] non ou difficilement préposables :

*déconcertant, drolatique, inouï, inusité*

Cette observation est en accord avec le fait que seuls les emplois de *curieux* préposé à son régissant nominal sont tous de type STIMULUS (sans que l'inverse soit exact).

(III) Ces trois valeurs sont manifestement liées à des schèmes morphologiques ou étymologiques :

— les adjectifs qui se terminent comme *curieux* par la syllabe [œ/œz] (à valeur de suffixe pour *désireux, soigneux* et *soucieux*, sans valeur morphologique en français moderne pour *anxieux* et *curieux*) dénotent une disposition à accomplir un type d'action<sup>17</sup> ;

— les adjectifs et substantifs qui se terminent par *-[at]eur* (en général à valeur de suffixe, sauf pour *inquisiteur* et *spectateur*) dénotent un comportement agentif :

*amateur, chercheur, collectionneur, flâneur, fouilleur, fouineur, fureteur, inquisiteur, investigateur, spectateur, voyeur*

— inversement les adjectifs dérivés de participes présents dénotent une propriété d'un STIMULUS. La construction verbale prend en général<sup>18</sup> un humain pour objet (ex. *qch déconcerte qn*), lequel peut être introduit par *pour* dans la construction adjectivale (ex. *qch est déconcertant pour qn*) :

*amusant, attachant, déconcertant, étonnant, extravagant, intéressant, piquant, plaisant, surprenant*

### 3.3. Visualisation de la polysémie de *curieux* à partir de la répartition des cliques de synonymes dans un espace sémantique multidimensionnel

Le logiciel VISUSYN permet de visualiser l'espace sémantique associé à une unité lexicale, en représentant par des points les cliques<sup>19</sup> du graphe de synonymie obtenu à partir du D.E.S. Dans le cas de *curieux*, le graphe de synonymie comporte 65 cliques. Une projection bidimensionnelle de l'espace sémantique obtenu est reproduite en Figure 3.1.

Comme on peut le constater, les sens de *curieux* se répartissent dans l'espace sémantique suivant trois branches<sup>20</sup> qui correspondent aux trois valeurs A, D, S, que nous avons définies ci-dessus. Ainsi la représentation géométrique, obtenue automatiquement, reflète assez fidèlement l'analyse linguistique.

<sup>17</sup> C'est déjà une propriété récurrente (mais non exclusive) du suffixe du latin tardif *-osus*, cf. *adfectuosus* > *affectueux*, *ambitiosus* > *ambitieux*, *bellicosus* > *belliqueux*, *captiosus* > *captieux*.

<sup>18</sup> A l'exception d' *extravaguer* intransitif.

<sup>19</sup> Une clique est un ensemble d'unités qui sont toutes synonymes entre elles deux à deux. En termes mathématiques plus précis, c'est un sous-graphe complet maximal du graphe de synonymie. Pour une présentation détaillée des algorithmes utilisés dans VISUSYN (calcul d'une distance entre cliques, projection bidimensionnelle de l'espace sémantique, etc.), voir Ploux et Victorri (1998).

<sup>20</sup> Le centre étant occupé par le sens nominal de *badaud*, que nous avons écarté dans la section précédente : cf. note 14

On peut noter en particulier deux caractéristiques intéressantes de cette représentation :

- Les cliques correspondent à des sens très précis de *curieux*. Notamment, certaines cliques représentent assez précisément les sens répertoriés par les dictionnaires historiques, comme on peut le constater sur le Tableau 3.1.

| données historiques  | cliques  |
|--|--|
| 12 <sup>e</sup> s. qualifie la personne qui s'occupe avec soin de qqch.          | { <i>attentif, curieux, soigneux, soucieux</i> }   |
| 16 <sup>e</sup> s. personne soucieuse de voir, de savoir, valeur morale négative | { <i>curieux, fouinard, fouineur, fureteur, indiscret</i> }<br>{ <i>curieux, fouineur, fureteur, indiscret, inquisiteur</i> }<br>{ <i>curieux, indiscret, voyeur</i> } |
| 17 <sup>e</sup> s. qui rassemble, recherche des objets rares ou précieux         | { <i>amateur, collectionneur, curieux</i> }<br>{ <i>chercheur, collectionneur, curieux, fouineur</i> }   |
| 18 <sup>e</sup> s. valeur méliorative (recherche désintéressée)                  | { <i>amateur, curieux, dilettante</i> }<br>{ <i>chercheur, curieux, fureteur, investigateur</i> }  |

**Tableau 3.1 :** Correspondances entre les définitions du dictionnaire historique et certaines cliques du graphe de synonymie

- On observe des changements graduels de sens le long des trois branches de la représentation : Pour la valeur S (branche en bas à gauche), le sens passe progressivement à partir du centre d'*intéressant* à *amusant* et finalement *bizarre*. De même, pour la valeur A (branche en bas à droite), il passe toujours à partir du centre des valeurs mélioratives (*amateur, collectionneur*) aux valeurs péjoratives (*fouineur, fouinard*).

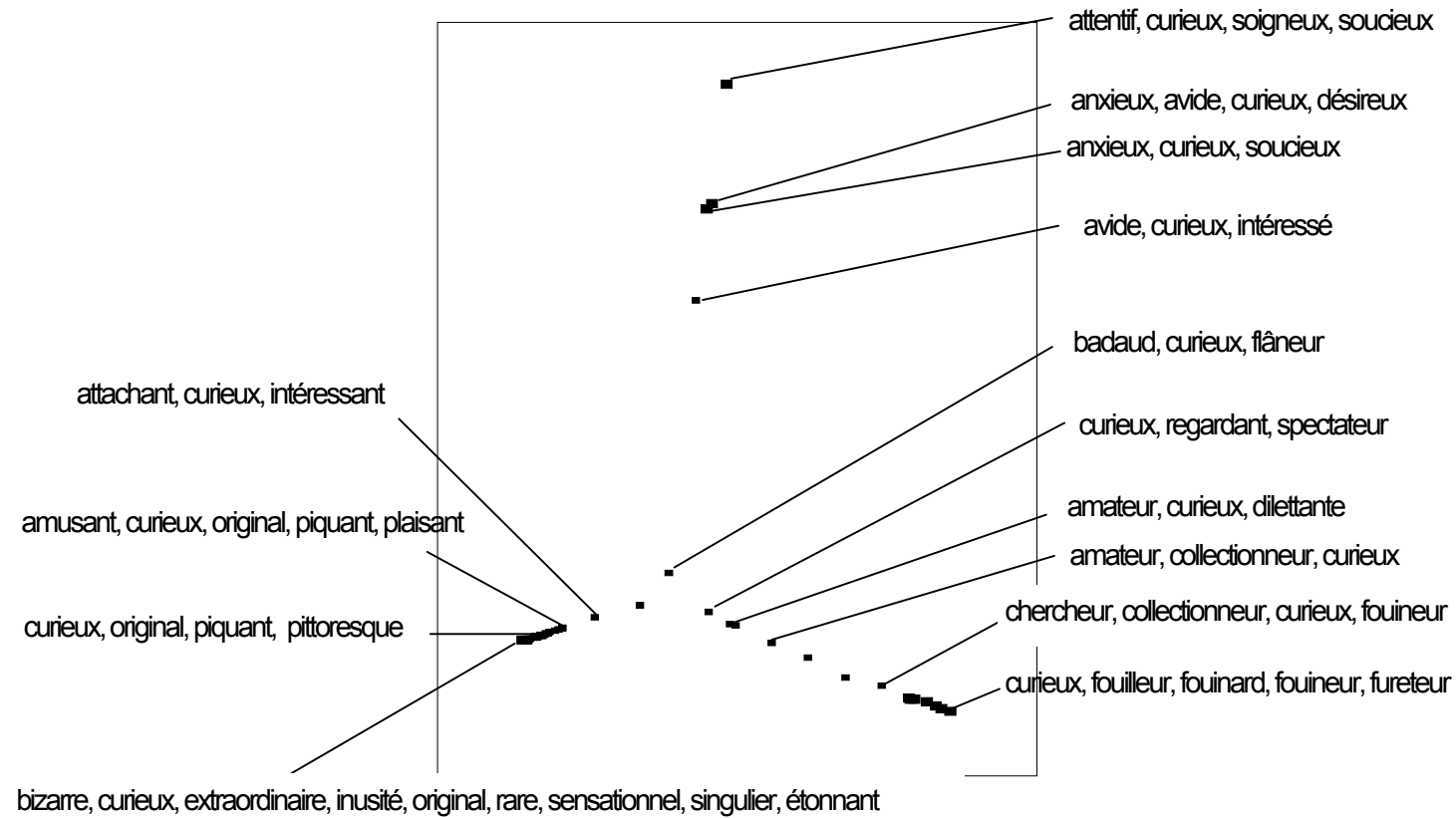
### 3.4. Analyse de l'influence du régissant nominal sur le sens de *curieux* à l'aide de la base FRANTEXT

Peut-on prédire la valeur de *curieux* dans un énoncé donné ? Comme nous l'avons vu, le problème est vite résolu quand l'adjectif est épithète préposée, puisque son antéposition n'est compatible qu'avec la valeur STIMULUS. En revanche, la difficulté reste entière pour les autres positions de l'adjectif (épithète postposée, attribut, etc.). Nous n'aborderons ici que le cas de l'épithète postposée.

La base de données textuelles FRANTEXT catégorisée permet d'effectuer des requêtes contextuelles spécifiant seulement la classe des constituants environnants. Pour la période 1850-1970<sup>21</sup>, elle livre 274 concordances du type N (Adv) *curieux*.

Dans un premier temps, nous nous sommes demandé dans quelle mesure la donnée du N régissant suffisait à désambiguïser l'adjectif. Nous avons mené cette étude à l'aide du logiciel VISUSYN, qui permet de calculer, à partir des données de FRANTEXT, une « fonction de compatibilité » d'un N donné avec les différents sens de *curieux*. Cette fonction est définie sur l'espace sémantique associé à l'adjectif.

<sup>21</sup> La base catégorisée s'arrête à la fin des années 60 (NB : ancienne version).



**Figure 3.1** : Représentation géométrique des cliques associées à *curieux*

La technique a été exposé dans la Section 1. Pour un N donné, on interroge FRANTEXT pour obtenir la fréquence de co-occurrence de cet N avec chacun des synonymes de *curieux*. A partir de ces données, on calcule ce que l'on appelle le « degré de co-occurrence » du N avec chacune des cliques du graphe de synonymie de *curieux*. Ce degré de co-occurrence est un bon indicateur de la compatibilité sémantique du N avec chacun des sens précis de l'adjectif (les points représentant les cliques sur l'espace sémantique). La fonction de compatibilité est alors obtenue en faisant, en chaque point de l'espace sémantique, une moyenne des degrés de co-occurrence pondérés par la distance de la clique au point considéré.

Pour *curieux*, cette technique a été testée sur une quarantaine de noms (les plus fréquents dans le corpus). Les résultats obtenus<sup>22</sup> font apparaître grossièrement deux catégories de N, en nombre à peu près égal :

- les N pour lesquels la fonction de compatibilité ne prend des valeurs positives que sur la zone stimulus de l'espace sémantique. Il s'agit de noms désignant des entités inanimées ou des événements qui ne peuvent effectivement être qualifiés par *curieux* qu'avec cette valeur, comme cas, chose, détail, fait, livre, mélange, objet, particularité, phénomène, scène, spectacle, etc. A titre d'exemple, on trouvera figure 3.2 la représentation bidimensionnelle (sous forme de courbes de niveau) de la fonction de compatibilité du nom spectacle.
- les N pour lesquels la fonction de compatibilité couvre les régions STIMULUS et AGENT (et/ou DISPOSITION<sup>23</sup>). Il s'agit principalement de noms désignant un animé (*animal, femme, fille, foule, homme, etc.*) ou une propriété, un comportement humain (*attention, esprit, œil, regard, etc.*). On trouvera figure 3.3 la fonction correspondant à *regard*.

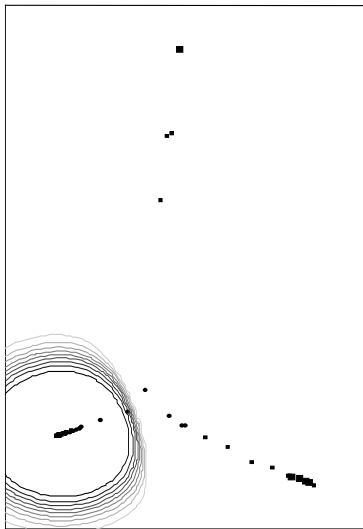
Ainsi, la donnée du nom régissant permet la désambiguïsation de l'adjectif *curieux* en position épithète postposée dans environ la moitié des cas, et on dispose d'un moyen automatique pour savoir si un nom peut à lui seul effectuer cette désambiguïsation.

Les fonctions de compatibilité sont représentées dans l'espace sémantique associé à *curieux* tel qu'il a été présenté au §3.3 : se reporter à la figure 3.1 pour la disposition des cliques)

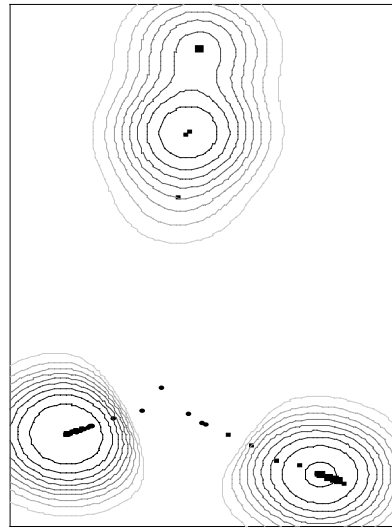
---

<sup>22</sup> Ces résultats sont encore préliminaires : un plus grand nombre d'expérimentations, portant sur divers adjectifs, et surtout avec des corpus plus larges, sera nécessaire pour valider complètement la méthode et ajuster plus finement les algorithmes utilisés.

<sup>23</sup> Dans la suite de cette section, nous regrouperons les valeurs AGENT et DISPOSITION, l'opposition principale étant entre STIMULUS d'une part, et AGENT/DISPOSITION de l'autre.



**Figure 3.2** : Fonction de compatibilité associée à *spectacle*



**Figure 3.3** : Fonction de compatibilité associée à *regard*

### 3.5. Prise en compte d'un contexte plus large pour la désambiguïisation des emplois postposés de *curieux*

Sur les 274 concordances du type N (Adv) *curieux* dans la base FRANTEXT catégorisée, on trouve 125 d'entre elles pour lesquelles le N permet la double lecture STIMULUS (SYN *intéressant*) vs. AGENT (SYN *intéressé*). Comme nous l'avons vu dans la section précédente, la seule spécification du régissant nominal est a priori inapte à lever l'ambiguïté dans ce cas. Notre propos est d'identifier la taille du contexte nécessaire à la sélection de l'une de ces deux lectures. Nous distinguerons cinq valeurs le long d'un continuum AGENT → STIMULUS :

|       |   |
|-------|---|
| A :   | lecture AGENT obligatoire ou très vraisemblable (incluant la lecture DISPOSITION) |
| ↓     |   |
| A? :  | lecture AGENT préférentielle  |
| ↓     |   |
| S/A : | lecture indéterminée en fonction du contexte disponible                           |
| ↓     |   |
| S? :  | lecture STIMULUS préférentielle   |
| ↓     |   |
| S :   | lecture STIMULUS obligatoire ou très vraisemblable                                |

Le tableau 3.2 classe les régissants nominaux par proportion décroissante de concordances avec *curieux* postposé propres à déclencher une lecture agentive obligatoire ou préférentielle [A/A?]. Compte tenu du faible nombre des concordances (à l'exception de *regard curieux*), les proportions ont une valeur toute relative. Dans le cas où la base fournit également pour ces régissants nominaux des concordances à adjectif antéposé, elles sont signalées dans la colonne de droite.

Cette disposition des concordances permet plusieurs observations :

- (I) Les 7 premiers régissants nominaux n'accueillent une qualification par *curieux* postposé qu'avec une valeur agentive (ou de disposition). Trois régissants : *filles*, *homme* et *visage* accueillent également *curieux* antéposé (obligatoirement avec la valeur STIMULUS).
- (II) Les 4 suivants {*regard*, *tête*, *esprit*, *femme*} présentent chacun une concordance indécidable en fonction du contexte fourni par la base FRANTEXT, ce qui signifie que les éléments manquent pour la constitution d'une scène mentale cohérente. *Regard*, *tête* et *esprit* présentent également des concordances avec *curieux* antéposé.
- (III) Enfin le troisième groupe est constitué de 7 régissants qui présentent tous au moins une concordance avec *curieux* postposé à valeur de stimulus. *Coup d'œil* et *étranger* présentent également chacun une concordance indécidable dans les limites du corpus.

|                           | adjectif postposé |           |          |          |           | total      | proportion  | adj.<br>antéposé |
|---------------------------|-------------------|-----------|----------|----------|-----------|------------|-------------|------------------|
|                           | A (D)             | A?        | S/A      | S?       | S         |            |             |                  |
| (I)                       |                   |           |          |          |           |            |             |                  |
| <i>air</i>                | 8                 | 1         |          |          |           | 9          | 1,00        |                  |
| <i>attention</i>          | 4                 | 1         |          |          |           | 5          | 1,00        |                  |
| <i>filles</i>             | 3                 |           |          |          |           | 3          | 1,00        | 1                |
| <i>foule</i>              | 3                 |           |          |          |           | 3          | 1,00        |                  |
| <i>homme</i>              | 1                 |           |          |          |           | 1          | 1,00        | 4                |
| <i>œil/yeux</i>           | 12                | 2         |          |          |           | 14         | 1,00        |                  |
| <i>visage</i>             | 3                 | 1         |          |          |           | 4          | 1,00        | 1                |
| (II)                      |                   |           |          |          |           |            |             |                  |
| <i>regard</i>             | 40                | 3         | 1        |          |           | 44         | 0,98        | 3                |
| <i>tête</i>               | 5                 | 4         | 1        |          |           | 10         | 0,90        | 1                |
| <i>esprit</i>             | 4                 |           | 1        |          |           | 5          | 0,80        | 2                |
| <i>femme</i>              | 1                 | 1         | 1        |          |           | 3          | 0,67        |                  |
| (III)                     |                   |           |          |          |           |            |             |                  |
| <i>figure</i>             | 1                 | 1         |          | 1        |           | 3          | 0,67        |                  |
| <i>coup d'œil</i>         | 2                 |           | 1        |          | 1         | 4          | 0,50        |                  |
| <i>expression</i>         | 1                 |           |          |          | 1         | 2          | 0,50        | 3                |
| <i>étranger</i>           | 1                 |           | 1        |          | 1         | 3          | 0,33        |                  |
| <i>animal</i>             | 1                 |           |          | 2        | 1         | 4          | 0,25        |                  |
| <i>bête</i>               |                   |           |          | 1        | 4         | 5          | 0,00        | 1                |
| <i>type</i> <sup>24</sup> |                   |           |          |          | 5         | 5          | 0,00        |                  |
| <b>total</b>              | <b>90</b>         | <b>14</b> | <b>6</b> | <b>4</b> | <b>13</b> | <b>127</b> | <b>0,83</b> | <b>16</b>        |

**Tableau 3.2 :** Classement des concordances [N *curieux*] de la base FRANTEXT catégorisée par proportion décroissante de la lecture agentive et avec mention des concordances [*curieux* N] pour chaque régissant nominal

Au total, outre les 16 concordances avec *curieux* antéposé qui imposent la lecture STIMULUS, 121 des 127 concordances avec *curieux* postposé sont décidables en faveur de la lecture AGENT (ou DISPOSITION) ou de la lecture STIMULUS, 104 pour la première lecture, 17 pour la seconde. Abstraction faite des six concordances indécidables, la lecture STIMULUS se partage donc presque à égalité entre des emplois antéposés [*curieux* N] et postposés [N *curieux*].

<sup>24</sup> Dans les cinq concordances, *type* est employé dans le sens d'espèce.

Si l'ordre [*curieux* N] a été introduit au 16<sup>e</sup> siècle pour lever des ambiguïtés, force est cependant de constater que cette possibilité n'est pas exploitée systématiquement et que dans de nombreux cas l'auteur se fie à l'aptitude du lecteur à constituer une scène mentale suffisamment cohérente pour lever l'ambiguïté. Sur quels éléments le lecteur peut-il prendre appui pour constituer cette scène ?

En premier lieu, la présence d'un complément prépositionnel prédicatif à l'infinitif permet de sélectionner la lecture DISPOSITION, par ex. *curieux de surprendre ...* en (1) :

(1) *Parmi tant de passants qui arrêtaient, sur cette image, des regards CURIEUX de surprendre les paroles que ces deux hommes pouvaient se dire (...)* [J. & J. Tharaud, *Dingley, l'illustré écrivain*, pp.133-13]

La lecture AGENT peut être sélectionnée à partir d'une modification adverbiale présupposant un agent, par ex. *ces regards ardemment curieux* en (2), *une foule de têtes cruellement curieuses* en (3) :

(2) *Il n'y avait aucune accusation, dans ces regards ardemment CURIEUX; et il croyait pourtant voir poindre le soupçon vague, ce doute que le plus petit fait parfois change en certitude.* [E. Zola, *La bête humaine*, p. 68]

(3) *Des juges, un greffier, des gendarmes, une foule de têtes cruellement CURIEUSES, il avait déjà vu cela une fois, autrefois, il y avait vingt-sept ans.* [V. Hugo, *Les misérables*, p. 322-323]

Le sémantisme de la construction régissante peut avoir le même effet. Ainsi par ex. en (4) *les regards curieux* est indécidable mais la construction à l'abri des regards curieux favorise la lecture agent parce que s'abriter de l'indiscrétion des regards fournit une scène mentale cohérente et s'abriter de l'étrangeté des regards une scène incohérente (cf. ?à l'abri de curieux regards).

(4) *En classe d'histoire, ce matin-là, il profita ce que, tout au fond de la salle, une place restait libre pour s'y asseoir à l'abri des regards CURIEUX, car il lui semblait que l'attention des uns et autres se dirigeait sur lui un peu plus que d'habitude.* [J. Green, *Moira*, p. 231]

De même la coordination entre *curieux* et un adjectif présupposant un agent comme *avide* déclenche la constitution d'une scène où cette valeur agentive s'étend à *curieux* comme en (5) :

(5) *Je m'étonne encore qu'elle ne sentît pas tous nos regards CURIEUX et avides collés pour ainsi dire à son bras nu.* [V. Larbaud, *Fermina Marquez*, p.24-25]

Inversement en (6), le sémantisme du verbe *figurer* et le contenu de la proposition relative *dont j'avais lu la description* favorisent la lecture STIMULUS, car les *animaux curieux* sont imaginaires et représentés seulement par le biais de leur description. Cependant cette description pourrait spécifier qu'ils manifestent de la curiosité et la sélection de la scène mentale n'est donc pas complètement décidable.

(6) *Ces autres pierres me figuraient, les unes des villes, les autres des animaux CURIEUX dont j'avais lu la description.* [P. Bourget, *Le disciple*, p. 76]

De même la sélection de la lecture STIMULUS est favorisée mais non fixée en (7) par la combinaison du regard extérieur porté sur l'*animal curieux* et de la spécification de celui-ci comme faisant partie de la ménagerie du prince. On regarde en priorité ce qui attire la curiosité et des animaux de ménagerie entrent dans cette catégorie, mais on ne peut pas exclure qu'une ménagerie comprenne des animaux intéressés par le monde extérieur.



- (7) *Quand il avait fini, on l'assomait de compliments, on le présentait de l'un à l'autre. Il pensait qu'on le regardait comme un animal CURIEUX, qui faisait partie de la ménagerie du prince, et que les éloges s'adressaient plus à son maître qu'à lui.* [R. Rolland, *Jean-Christophe, Le matin*, pp.114-115]

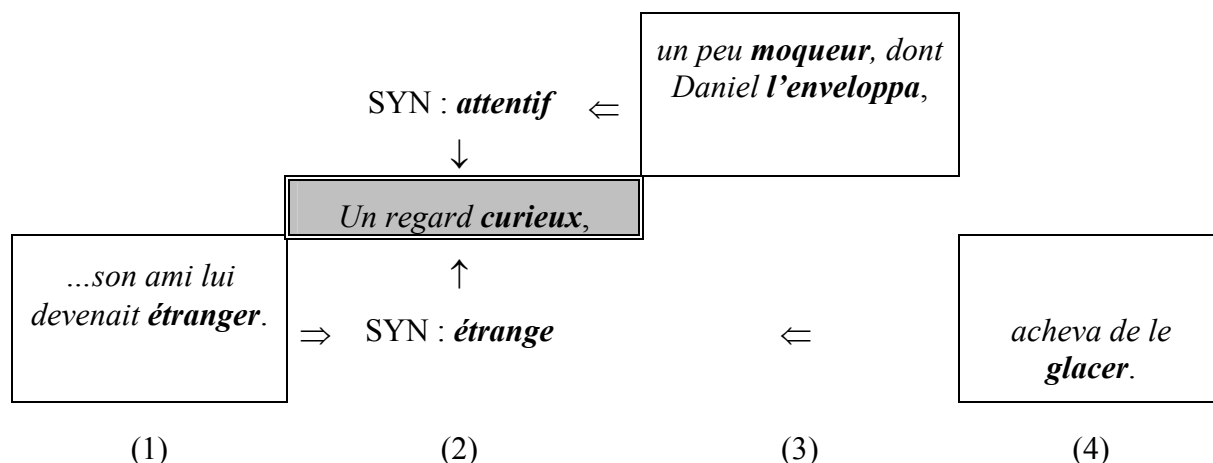
La concordance (8) est particulièrement intéressante, car elle met en jeu quatre adjectifs dépendants de *regard* et coordonnés deux par deux. Les constituants de la seconde paire sont antinomiques : *implacable* s'oppose à *indulgent*. On peut donc supposer que la même antinomie oppose ceux de la première paire. Or *curieux* ne peut s'opposer à *indifférent* qu'avec une valeur agentive (*bizarre / étrange* n'entre pas en opposition avec *indifférent*). En conséquence, la sélection de la lecture agentive présuppose la reconnaissance d'un jeu rhétorique.

- (8) *Je montai dans la chambre, toujours accompagnée du regard infiniment CURIEUX et indifférent, implacable et indulgent, de la supérieure.* [P. Drieu La Rochelle, *Réveuse bourgeoise*, pp. 342-343]

Pour conclure cette section, il nous faut encore présenter l'une au moins des six concordances indécidables au vu du contexte fourni par FRANTEXT. En (9) *un regard curieux* peut être compris comme un regard étrange sous l'effet du contexte antérieur, *son ami lui devenait étranger*. En revanche des éléments du contexte immédiatement postérieur, *moqueur* et *envelopper* [qn d'un regard] vont dans le sens de l'intérêt de Daniel pour Jacques, mais le contexte ultérieur *acheva de le glacer* favorise à nouveau le caractère étrange du regard de Daniel.

- (9) *Jacques eut une impression pénible, comme s'il eût essuyé une offense. De minute en minute son ami lui devenait étranger. Un regard CURIEUX, un peu moqueur, dont Daniel l'enveloppa, acheva de le glacer. - " mais toi ? " questionna Daniel, en se rapprochant. " d'après ta lettre, toi aussi, tu..."* [R. Martin du Gard, *Les Thibaud. Le pénitencier*, p. 792]

On peut représenter ainsi le va-et-vient entre les deux lectures STIMULUS (à partir de 1 et 4, SYN : *un regard étrange*) et AGENT à partir de (3, SYN : *un regard attentif*) :



### 3.6. Bilan

Le *D.E.S.* permet une analyse non componentielle de la polysémie adjectivale en fournissant des visualisations de l'espace sémantique d'un adjectif polysémique tel que *curieux* dont chaque point est occupé par une clique de synonymes. Lorsque la liste des synonymes fournis par la nomenclature du *D.E.S.* est abondante, les cliques se disposent comme les tuiles d'un toit et on s'aperçoit intuitivement que pour chaque clique l'adjectif étudié et la totalité ou la plupart de ses synonymes dans cette clique peuvent être régis par un même nom et que cela vaut à un degré plus ou moins élevé pour une suite de cliques 'tuilées'. Cette observation intuitive peut être confirmée en calculant la « fonction de compatibilité » entre l'espace sémantique d'un adjectif et celui d'un nom susceptible de le régir. La réaction nominale a donc — comme on pouvait le supposer — une fonction de sélection d'une zone plus ou moins étroite du champ sémasiologique de l'adjectif. Toutefois cette fonction échoue fréquemment à réduire radicalement le champ interprétatif, c'est-à-dire à permettre la constitution d'une scène mentale précise et cohérente et on peut supposer que l'interprétant (angl. *comprehender*) précise la scène — dans les limites de ses besoins immédiats de compréhension qui peuvent éventuellement s'accommoder d'une scène floue mais non d'une scène incohérente — en tirant parti de l'effet sélectif du sens de différents constituants du cotexte (adverbes, adjectifs coordonnés, constructions à effet rhétorique, etc.). Dans de rares cas, la convocation du cotexte ne permet pas de décider en faveur de lectures mutuellement incompatibles. Dans un discours informatif, on peut sans doute y voir une défaillance de l'auteur, mais la superposition des lectures (voir dans l'extrait 9 *un regard curieux*) peut au contraire être recherchée dans le discours littéraire.

Dans la section 4, nous porterons notre attention sur la possibilité offerte par le *D.E.S.* combiné avec la base FRANTEXT catégorisée, d'analyser l'intersection entre l'espace sémantique de deux adjectifs de très haute fréquence, en l'occurrence *gras* et *gros*, qui bien que fréquemment coordonnables (*une volaille bien grosse et grasse*), sont également occasionnellement synonymes. Le domaine d'intersection est celui où ces deux adjectifs ont des synonymes en commun. Mais les ont-ils dans les mêmes types de cotexte nominal ?

## Section 4. Deux adjectifs occasionnellement synonymes : *gros* et *gras* en cotexte nominal

Au §4.1 nous donnerons une première image superficielle de la polysémie des deux adjectifs, telle qu'elle ressort de l'inventaire des synonymes délivré par le *D.E.S.* 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 co-textuelles (4.3.2). 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 (§4.3). En conclusion (§4.4) nous évoquerons les problèmes originaux auxquels nous avons confrontés l'étude combinée des deux champs sémasiologiques de *gros* et de *gras*, par rapport à ceux que nous avons rencontrés précédemment dans l'étude de la polysémie de *curieux* (Section 3).

### 4.1. *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 :

- (I) le relevé de co-occurrences régulières
- (II) la substitution de synonymes et d'antonymes propres à conserver ou inverser le sens du terme vedette dans le type de contexte ainsi identifié
- (III) et la limitation à un registre stylistique et/ou à un domaine de spécialité particuliers.

Nous allons nous arrêter dans un premier temps à l'inventaire brut des relations synonymiques du *D.E.S.* constitué à partir des listes de renvois synonymiques de sept dictionnaires compilées par l'INaLF pour l'élaboration du *Trésor de la Langue Française* (cf. Section 1) avant de passer à l'exploitation des cliques de synonymes. Le *D.E.S.* 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, boursoufflé, caniculaire, charnu, cher, colossal, commun, conséquent, considérable, coquet, corpulent, courtaud, développé, dodu, élevé, éminent, empâté, en bloc, enflé, énorme, épais, épanoui, essentiel, étouffant, excessif, exorbité, extraordinaire, fabuleux, famélique, fécond, fertile, formidable, fort, gênant, gidouillard, globalement, globuleux, gonflé, grand, gras, grave, gravos, gros bonnet, grossi, grossier, haut, hypertrophié, immense, important, imposant, influent, insupportable, intense, joli, joufflu, large, 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, 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, boursoufflé, 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, lesté, 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*

A proprement parler, ces listes ne sont 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]  $\approx$  *une barrique, un tonneau, etc.*}, soit des êtres humains {*un gros*[+humain]  $\approx$  *un nabab, une notabilité, un notable, etc.*}. Cela vaut à un moindre titre pour *gras* {*le gras*  $\approx$  *la graisse, le lard*}.
- Ensuite on remarque que certains synonymes de *gros* ou *gras* ne le sont que dans des cotextes 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, lesté, 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 4.1) les conditions de synonymie entre {*gras, abondant, bouffi, corpulent, épais, graisseux, grivois, salé*} dans les contextes nominaux {*papier, herbe, personne, plaisanterie*} :

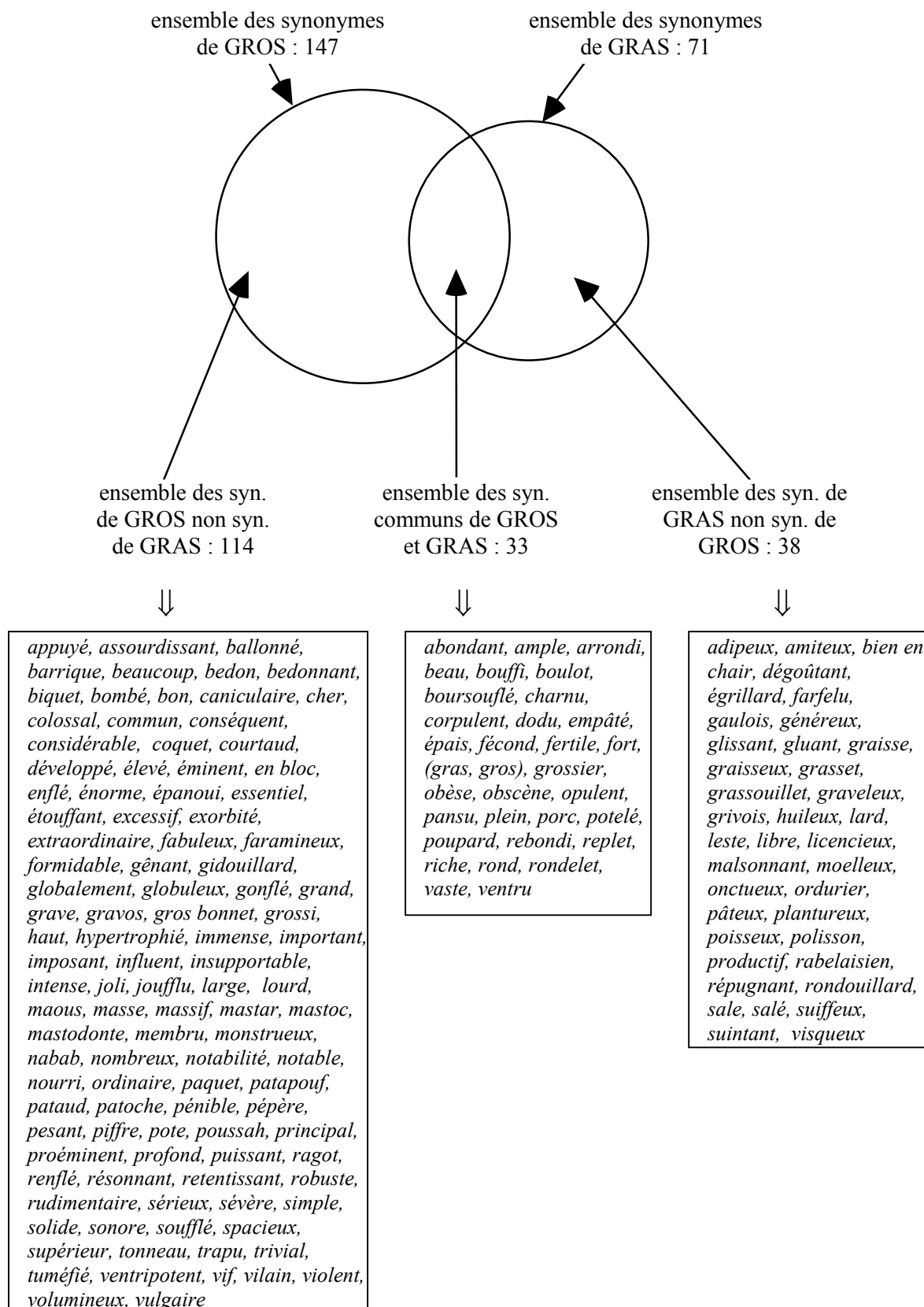
| N <i>gras(se)</i> | <i>herbe</i> | <i>papier</i> | <i>personne</i> | <i>plaisanterie</i> |
|-------------------|--------------|---------------|-----------------|---------------------|
| <i>abondant</i>   | SYN          | ≠             | *               | *                   |
| <i>bouffi</i>     | *            | *             | SYN             | *                   |
| <i>corpulent</i>  | *            | *             | SYN             | *                   |
| <i>épais</i>      | SYN          | ≠             | ±SYN            | ±SYN                |
| <i>graisseux</i>  | ≠            | SYN           | ≠               | *                   |
| <i>grivois</i>    | *            | *             | ≠               | SYN                 |
| <i>salé</i>       | ≠            | ?             | *               | SYN                 |

**Tableau 4.1** : 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 donc 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 lui est pas 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 4.1).



**Figure 4.1.** L'espace sémantique commun aux deux listes des synonymes de *gros* et *gras* fournies par le DES

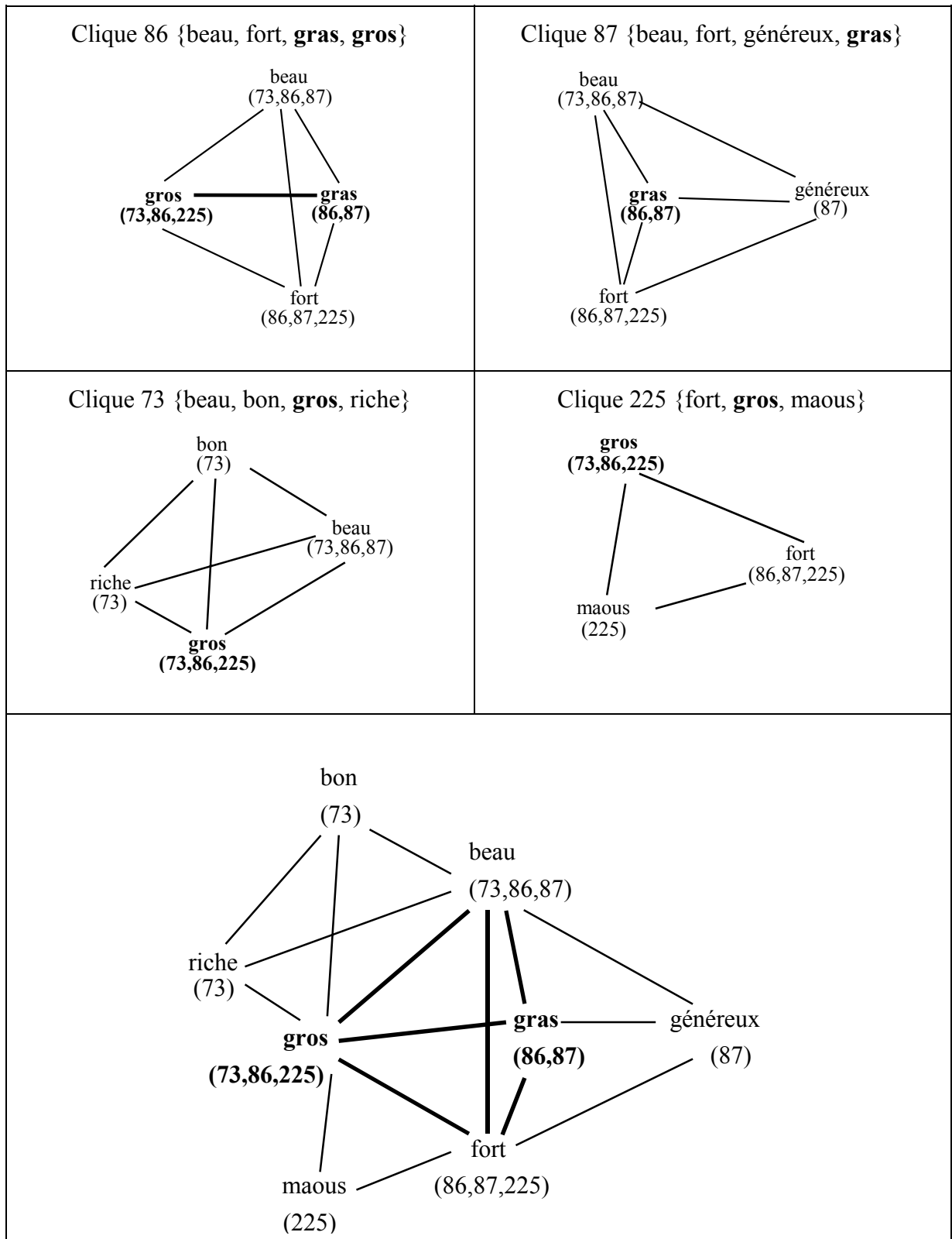
Mais parler d'un espace sémantique commun laisse entendre que cet espace soit unifié. Or force est de constater que si la plupart des adjectifs de cet espace peuvent qualifier des personnes ou des parties du corps où le caractère *gros* est supposé en rapport avec un excès de graisse, d'autres s'appliquent à des comportements ou des actes {*grossier, obscène*} et *abondant* désigne l'effectif élevé d'un collectif (*herbe, cheveu*)<sup>25</sup>. 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*.

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

Nous appliquons ici à nouveau la méthode exposée en section 1. Afin de clarifier le phénomène déjà mentionné de 'tuilage' entre les cliques de synonymes, 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 4.2). 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 *génereux*. Aucune autre clique ne comporte de lien entre *gros* et *génereux*, il n'y a donc pas de clique associant *génereux* 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 *génereux*, adjectifs pourtant manifestement synonymes. Nous nous sommes imposé de ne garder que les cliques contenant au moins une des deux vedettes, si bien que nous ferons ici abstraction des quelque 35 cliques qui incluent à la fois *bon* et *génereux*, mais ni *gros* ni *gras*.

---

<sup>25</sup> *Herbe* et *cheveu* ne peuvent être *abondants* qu'en tant que collectifs.



**Figure 4.2.** Représentation éclatée et combinée des quatre cliques 73, 86, 87 et 225

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, de dimension 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, nous avons pu combiner les plans qui dissocient certains sens et en obtenir un nouveau qui offre une vue plus synthétique évitant les chevauchements trompeurs ; nous avons déjà employé cette opération dans des études antérieures<sup>26</sup>. La lecture d'une telle représentation devient alors plus aisée.

On voit ainsi, sur la figure 4.3, 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**<sup>27</sup>, 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 3.3 sur l'espace des synonymes de *gras* (et accessoirement de *gros*) et dans la figure 3.4 sur celui des synonymes de *gros* (et accessoirement de *gras*). Les deux figures 3.3 et 3.4 ont en commun la représentation de l'espace des deux adjectifs *gras* (à gauche) et *gros* (à droite). Sur la figure 3.3 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*<sup>28</sup>. 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 (*Paul est potelé*, ? *rond*, ?? *rebondi*, ?? *arrondi*).

---

<sup>26</sup> cf. Manguin J.L. & Victorri B. (1999)

<sup>27</sup> Les composantes ici numérotées 25 et 26 correspondent à des combinaisons linéaires des premières composantes données par l'analyse ; leur utilisation ne sert qu'à améliorer la lisibilité de la représentation.

<sup>28</sup> 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.

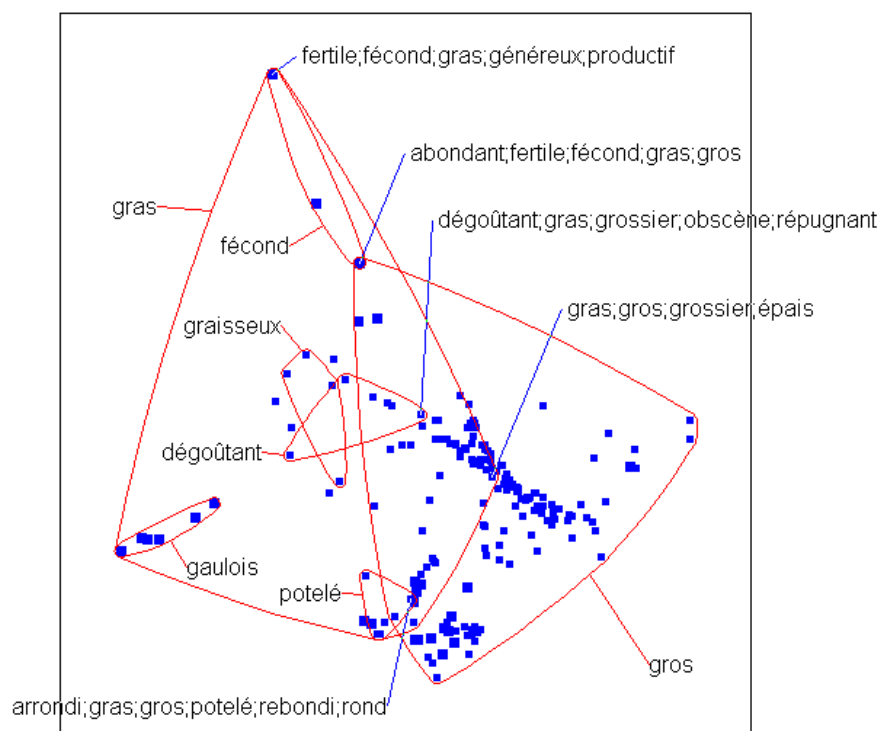


Figure 4.3 : Espace sémantique de *gros*, *gras*, de 3 synonymes de *gras* et de 2 synonymes communs aux deux.

Sur la figure 4.4 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*, *ventripotent*, etc.), soit transitoire et qualifiant une partie du corps (espace de *tuméfié* associé à *ballonné*, *boursoufflé*, *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 *griivois*, *égrillard* et *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).

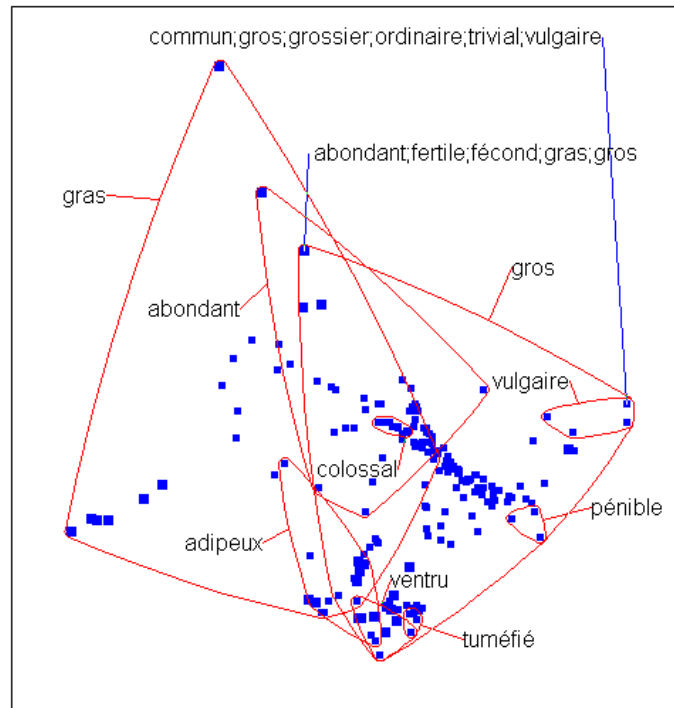


Figure 4.4 : Espace sémantique de *gros*, *gras*, de 4 synonymes de *gros* et de 2 synonymes communs aux deux.

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

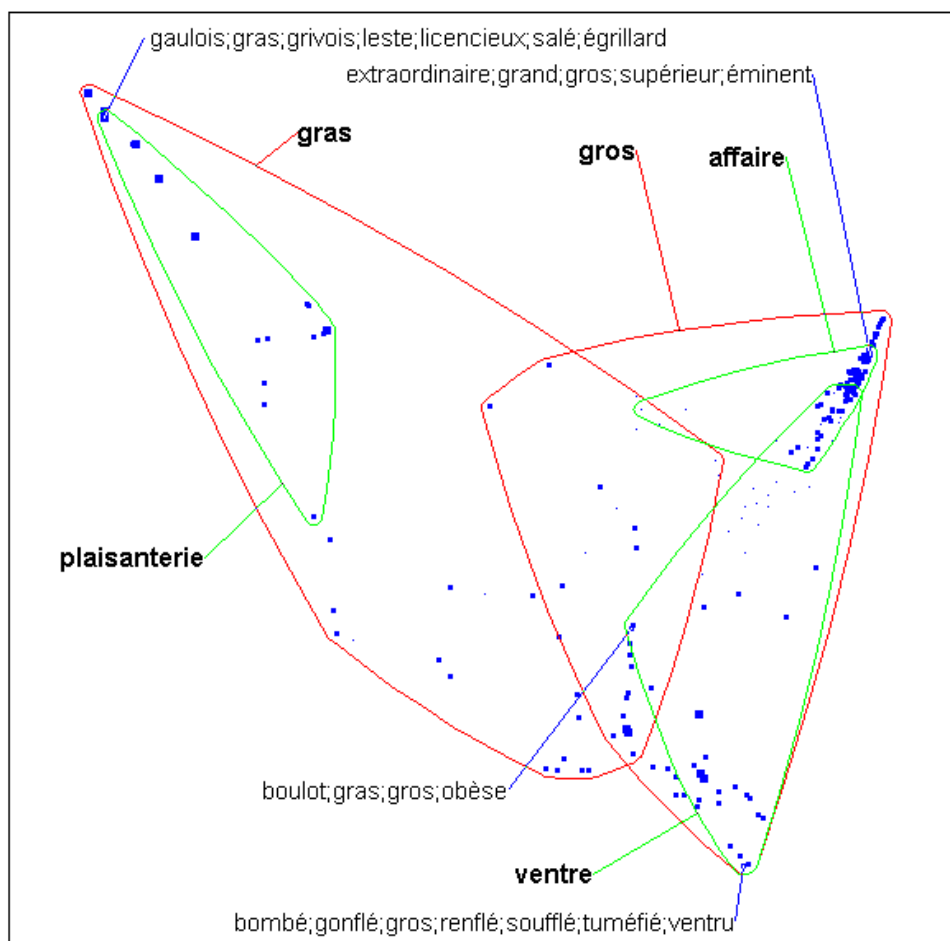
L'application de la méthode d'exploitation combinée du *D.E.S.* et de la base FRANTEXT catégorisée exposée en section 1 nous a permis de regrouper quelque 20 000 co-occurrences en 1842 paires (nom, adjectif) ; le Tableau 3.2 donne les paires obtenues avec les trois noms *affaire*, *plaisanterie* et *ventre* 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. | Reg. Nom.           | Adj.              | Nb occ.         | Reg. Nom.         | Adj.             | Nb occ.      |   |
|----------------------|-----------------------|---------|---------------------|-------------------|-----------------|-------------------|------------------|--------------|---|
| <b>affaire</b>       | <i>grande</i>         | 51      |                     | <i>énorme</i>     | 7               |                   | <b>gros</b>      | 69           |   |
|                      | <b>grosse</b>         | 48      |                     | <i>lourde</i>     | 6               |                   | <i>énorme</i>    | 23           |   |
|                      | <i>importante</i>     | 35      |                     | <i>obscène</i>    | 5               |                   | <i>ballonné</i>  | 13           |   |
|                      | <i>sérieuse</i>       | 34      |                     | <i>grossière</i>  | 4               |                   | <i>gonflé</i>    | 10           |   |
|                      | <i>grave</i>          | 21      |                     |                   | <b>grasse</b>   | 3                 |                  | <i>rond</i>  | 9 |
|                      | <i>sale</i>           | 16      |                     |                   | <i>vulgaire</i> | 2                 |                  | <i>enflé</i> | 6 |
|                      | <i>simple</i>         | 14      | <b>plaisanterie</b> | <i>colossale</i>  | 1               | <b>ventre</b>     | <i>large</i>     | 6            |   |
|                      | <i>considérable</i>   | 10      |                     | <i>dégoûtante</i> | 1               |                   | <i>arrondi</i>   | 5            |   |
|                      | <i>colossale</i>      | 5       |                     | <i>égrillarde</i> | 1               |                   | <i>bombé</i>     | 4            |   |
|                      | <i>vilaine</i>        | 5       |                     | <i>grande</i>     | 1               |                   | <i>lourd</i>     | 4            |   |
|                      | <i>pénible</i>        | 4       |                     | <i>massive</i>    | 1               |                   | <i>rebondi</i>   | 3            |   |
|                      | <i>énorme</i>         | 3       |                     | <i>ordinaire</i>  | 1               |                   | <i>rondelet</i>  | 3            |   |
|                      | <i>principale</i>     | 3       |                     | <i>sale</i>       | 1               |                   | <i>obscène</i>   | 2            |   |
|                      | <i>vaste</i>          | 3       |                     | <i>salée</i>      | 1               |                   | <i>puissant</i>  | 2            |   |
|                      | <i>immense</i>        | 2       |                     |                   |                 |                   | <i>empâté</i>    | 1            |   |
|                      | <i>lourde</i>         | 2       |                     |                   |                 |                   | <i>épanoui</i>   | 1            |   |
|                      | <i>extraordinaire</i> | 1       |                     |                   |                 |                   | <i>fécond</i>    | 1            |   |
|                      | <i>formidable</i>     | 1       |                     |                   |                 |                   | <i>haut</i>      | 1            |   |
|                      | <i>haute</i>          | 1       |                     |                   |                 |                   | <i>immense</i>   | 1            |   |
|                      | <i>large</i>          | 1       |                     |                   |                 |                   | <i>important</i> | 1            |   |
| <i>ordinaire</i>     | 1                     |         |                     |                   |                 | <i>monstrueux</i> | 1                |              |   |
| <i>retentissante</i> | 1                     |         |                     |                   |                 | <i>poisseux</i>   | 1                |              |   |
| <i>supérieure</i>    | 1                     |         |                     |                   |                 | <i>renflé</i>     | 1                |              |   |
|                      |                       |         |                     |                   |                 | <i>solide</i>     | 1                |              |   |
|                      |                       |         |                     |                   |                 | <i>spacieux</i>   | 1                |              |   |
|                      |                       |         |                     |                   |                 | <i>trapu</i>      | 1                |              |   |
|                      |                       |         |                     |                   |                 | <i>visqueux</i>   | 1                |              |   |

**Tableau 4.2** : Co-occurrences 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. Le degré de co-occurrence de chaque nom avec chacune des cliques du graphe de la synonymie est égal à la valeur de la « fonction de compatibilité » du nom pour chaque point de l'espace où se trouve une clique, autrement dit pour les différents sens présents sur la représentation (Figure 3.5). Pour des raisons de visibilité, il n'est pas possible de faire figurer sur le graphique les fonctions de plusieurs noms simultanément.

VisuSyn : *grosgras* (132 unités, 198 cliques, 88 cooccurrences) - comp. 1 et 2 , seuil 86 %



**Figure 4.5** : Représentation de la fonction de compatibilité entre les 198 cliques de *gras* et *gros* et trois régissants nominaux : *affaire*, *plaisanterie* et *ventre*<sup>29</sup>

Sur la figure 4.5, 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 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 (dans le calcul de la fonction de compatibilité, chaque adjectif et chaque nom sont dotés d'un poids qui dépend de leur fréquence relative d'apparition dans le corpus).

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 représentée sur un même graphique. La

<sup>29</sup> 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.

comparaison dans le tableau 4.3 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<sup>30</sup> 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.

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

**Tableau 4.3** : 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 est proposée en Section 2 pour l'espace sémantique de *curieux* (Figures 2.2 et 2.3) et sous une forme tridimensionnelle en Section 4 pour celui de *sec*. 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.

#### 4.4. Bilan

L'étude de l'espace sémantique de *curieux* (Section 3) nous a posé un problème spécifique, celui de sa désambiguïsation en position post-nominale. Pour *gros* et *gras*, rien de tel; l'un et l'autre peuvent apparaître dans les deux positions<sup>31</sup> mais sans incidence sémantique notable (contrairement par exemple à la paire distinctive *un homme grand* vs. *un grand*

<sup>30</sup> L'intuition ne devrait jouer aucun rôle dans notre méthode, mais une séquence intuitivement fréquente telle que *plaisanterie grivoise* se révèle absente de la base Frantext catégorisée dont nous nous sommes contentés jusqu'à présent.

<sup>31</sup> *Gras* se rencontre cependant beaucoup moins fréquemment que *gros* en antéposition (24 occurrences de *grasse matinée* et 7 de *gras pâturages*).

*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 dos* / [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, cœur, dame, doigt, dos, face, femme, figure, fille, garçon, goutte, homme, joue, lèvres, main, monsieur, mot, rire, soulier, trait, voix.*

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. *une 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 présence sous-jacente 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.

## 5. Conclusion

Le modèle de polysémie présenté ici présente trois traits principaux intéressants

- (I) Des relations de proximité sémantique entre différents sens d'un item polysémique sont pleinement prises en compte dans la représentation géométrique de l'espace sémantique.
- (II) Les différents types d'interprétation (sens précis, indéterminations, ambiguïtés) sont distingués par la forme et la taille de la région de l'espace sémantique associée à l'occurrence en question de l'item polysémique
- (III) Le mécanisme d'interprétation est considéré comme un processus dynamique de construction du sens qui est le même quelque soit le type d'interprétation en cause.

Comme nous l'avons montré, le modèle peut être implémenté en recourant à des ressources électroniques (dictionnaires et corpus). L'espace sémantique associé à un item polysémique est construit par une procédure automatique. Concernant l'automatisation du calcul du sens d'un item polysémique dans un contexte donné, nous avons dégagé des résultats prometteurs en calculant automatiquement le sens d'items polysémiques dans un contexte donné, mais notre travail se poursuit : nous avons à réaliser plus d'expériences afin de valider l'algorithme utilisé.

Il faut souligner que le même modèle peut être utilisé pour comparer des items polysémiques de deux langues différentes, comme l'ont montré Ploux et son équipe (cf. Ploux & Hi 2003), qui ont conçu un « atlas sémantique » multilingue. Ils ont construit ce qu'on peut appeler des « graphes de synonymie bilingue » en mélangeant les unités linguistiques de deux langues (à l'aide de deux dictionnaires monolingues de synonymes et d'un dictionnaire bilingue). Les cliques d'un tel graphe contiennent des items des deux langues. De ce fait, dans l'espace sémantique d'un item de la langue source on peut représenter la région associée à toute traduction possible de cet item dans la langue cible. Dans ce cadre, notre travail pourrait être exploité pour concevoir un système de T.A.O. puisque le calcul de la fonction potentielle associée à un contexte donné pourrait fournir automatiquement les traductions convenables de l'item polysémique dans ce contexte. En dehors de cette application pratique, de tels atlas sémantiques multilingues peuvent être utilisés pour étudier les différentes configurations dans le même espace géométrique. Ainsi ce modèle de polysémie devrait se révéler le plus intéressant dans la comparaison de la polysémie entre langues, puisqu'il peut prendre en compte la diversité lexicale aussi bien entre langues que dans une même langue.

## Bibliographie

- Bailly René (1947), *Dictionnaire des synonymes*, Paris, Larousse.
- Bat-Zeev Schyldkrot, H (1997) , « Synonymie et polysémie, le cas de CURIEUX comme parcours sémantique d'un mot ». *Langages* 128, 113-125
- Bénac Henri (1956), *Dictionnaire des synonymes*, Paris, Hachette.
- Berge C. (1970): *Graphes et hypergraphes*, Paris, Dunod.
- Bertaud du Chazaud Henri (1989), *Dictionnaire des synonymes*, Paris, Editions le Robert.
- Bourigault D., Jacquemin Ch., L'Homme M.-C. (2001): *Recent Advances in Computational Terminology*, Amsterdam, John Benjamins.
- Darmesteter, A. (1887) : *La vie des mots étudiée dans leurs significations*, Paris, Delagrave.



- Fellbaum C. (1998): *Wordnet: an Electronic Lexical Database*, Cambridge, MIT Press.
- François, J. & Manguin, J.L. (à paraître) : « La polysémie adjectivale entre synonymie et sélection contextuelle : le cas de *propre* ». In : H. Bouillon (ed.) Volume d'hommage à Jacques LEROT à l'occasion de son 65e anniversaire, Louvain, Presses universitaires.
- François J., Victorri B., Manguin J.L. (2002): Polysémie adjectivale et synonymie: l'éventail des sens de *curieux*, in Soutet O. (ed.) *La polysémie*, Paris, Presses de l'Université de la Sorbonne.
- Guilbert, L et alii (1971). *Grand Larousse de la Langue Française*, Paris, Larousse.
- Guizot François (1864), *Dictionnaire Universel des synonymes de la Langue Française*, Paris, Didier, (7è édition).
- Habert B., Nazarenko A., Salem, A. (1997): *Les linguistiques de corpus*, Paris, A. Colin.
- Ide N., Véronis J. (1990): Word Sense Disambiguation with very large neural networks extracted from machine readable dictionaries., *Proceedings of the 14th International Conference on Computational Linguistics (COLING'90)*
- Kahlmann A. (1975), *Traitement automatique d'un dictionnaire de synonymes*, Thèse de l'Université de Stockholm.
- Kleiber G. (1990): *La sémantique du prototype : Catégories et sens lexical*, Paris, P.U.F.
- Lafaye Pierre-Benjamin (1858), *Dictionnaire des synonymes de la Langue Française*, Paris, Hachette.
- Lakoff G. (1987): *Women, Fire and Dangerous Things*, University of Chicago Press.
- 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., François J., Victorri B. (2003): Polysémie adjectivale et rection nominale: quand *gros* et *gras* sont synonymes, in François J. (ed.), *L'adjectif en français et à travers les langues*, Presses Universitaires de Caen.
- Manguin J.L., Victorri B. (1999): Représentation géométrique d'un paradigme lexical, *Actes de la 8<sup>ème</sup> conférence TALN*, vol. 1:363-368.
- Niklas-Salminen (1997) : *La lexicologie*, Paris, Armand Colin.
- Nyckees, V. (1998) : *La Sémantique*, Paris, Belin.
- Pierrel J.M. (ed.) (2000) *Ingénierie des langues*, Paris, Hermès.
- Ploux S., Ji H. (2003), A Model for Matching Semantic Maps Between Languages (French / English, English / French ), *Computational Linguistics*, 29/2:155-178, Cambridge, MIT Press.
- Ploux S., Victorri B. (1998) : Construction d'espaces sémantiques à l'aide de dictionnaires de synonymes, *Traitement automatique des langues*, 39/1:161-182.
- Pustejovsky J. (1995): *The generative lexicon*, Cambridge, MIT Press.
- Quemada B. (ed.) (1992): *Trésor de la langue française*, tome XV, Paris, CNRS/Gallimard.
- Rey, A (éd. 1992). *Dictionnaire historique de la langue française*. Paris, Editions Le Robert.
- Rey, A et alii (1985). *Le Grand Robert. Dictionnaire de la Langue Française*, Paris, Editions Le Robert.

- Reingold E.M., Nievergelt J., Deo N. (1977): *Combinatorial Algorithms, Theory and Practice*, Prentice-Hall.
- Sadock J.M. (1986): Vagueness as a vague concept, *Quaderni di Semantica*, 7-2.
- Touratier, Ch. (2000), *La sémantique*. Paris, Armand Colin.
- Victorri B. (2002): Espaces sémantiques et représentation du sens, in *Textualités et nouvelles technologies, éc/artS*, 3.
- Victorri B., Fuchs C. (1996): *La polysémie – Construction dynamique du sens*, Paris, Hermès.
- Warnesson I. (1985): Applied Linguistics : Optimization of Semantic Relations by Data Aggregation Techniques, *Journal of Applied Stochastic Models and Data Analysis*, vol.1/2:121-143.
- Wittgenstein L. (1961): *Investigations philosophiques*, Paris, Gallimard.

Le Centre de Recherches Interlangues sur la Signification en Contexte (CRISCO) contribue au débat linguistique à l'aide de trois outils éditoriaux :

## I

### **Syntaxe & Sémantique**

Revue de linguistique créée en 2000 aux Presses Universitaires de Caen

**Responsables de la revue** : J. FRANÇOIS & N. LE QUERLER

**Comité de rédaction** : C. GUIMIER, A. LACHERET, R. RENAULT, J. FRANÇOIS, N. LE QUERLER

**Comité scientifique** : P. BLUMENTHAL (Cologne), A. BORILLO (Toulouse-le-Mirail), M. CHAROLLES (Paris 3), B. COMBETTES (Nancy), J.P. DESCLES (Paris 4), D. GEERAERTS (Leuven), B.N. GRUNIG (Paris 8), ), P. HIRSCHBUHLER (Ottawa), G. KLEIBER (Strasbourg 2), P. KOCH (Tübingen), B. LAMIROY (Leuven), A. LEMARECHAL (Paris 4), R. MARTIN (INALF), S. MEJRI (Tunis), C. VET (Groningen), B. VICTORRI (CNRS).

- n°1 **Connecteurs et marqueurs de connexions**, sous la direction de Claude GUIMIER (mars 2000)
- n°2 **Sémantique du lexique verbal**, sous la direction de Françoise CORDIER, Jacques FRANÇOIS et Bernard VICTORRI (avril 2001)
- n°3 **Les grammaires du français et les "mots outils"**, sous la direction de Gudrun LEDEGEN et Nathalie ROSSI (mars 2002)
- n°4 **La valence vue d'Allemagne**, sous la direction de Peter BLUMENTHAL et Peter KOCH (juin 2002)
- n°5 **Polysémie et polylexicalité**, sous la direction de Salah MEJRI (à paraître au 2<sup>e</sup> semestre 2003)

## II

### **Bibliothèque de Syntaxe & Sémantique**

collection de linguistique créée en 2002 aux Presses Universitaires de Caen

**Traits d'union**, sous la direction de N. LE QUERLER & G. KLEIBER [Actes des *Journées de Linguistique Rhénane* de novembre 2000, organisées par SCOLIA et le CERLICO] (mai 2002)

**L'adjectif en français et à travers les langues**, sous la direction de J. FRANÇOIS [Actes du colloque de Caen, 28-30 juin 2001] (à paraître au 2<sup>e</sup> semestre 2003)

### III

#### **Cahiers du CRISCO**

Cahiers de recherche linguistique diffusés par le CRISCO depuis 2000. Les **Cahiers du CRISCO** sont distribués gratuitement. Il est demandé une participation aux frais d'envoi.

- n°1 Jacques FRANÇOIS, *Sur la grammaire de van Valin* (septembre 2000)
- n°2 Véronique LENEPEU, *Adjectifs et adverbes : une corrélation syntactico-sémantique* (janvier 2001)
- n°3 Jacques FRANÇOIS & Hansjakob SEILER, *Gespräche über Hansjakob Seilers Universalienforschung* (mai 2001, version originale en allemand)
- n°4 Denis RAMASSE, *L'intonation des phrases présentant un détachement à gauche en français : une première approche* (juin 2001)
- n°5 Heike BALDAUF, *Voyage au bout de la langue : quelques remarques sur des énoncés minimaux* (mars 2002)
- n°6 Jacques FRANÇOIS (avec la participation de Hansjakob SEILER, *La faculté de langage et les universaux : perspectives fonctionnalistes actuelles* (mars 2002)
- n° 7 Catherine CAMUGLI GALLARDO, *La langue des jeunes en Italie — Guide à une enquête de terrain* (mars 2002)
- n° 8 Jacques FRANÇOIS, *La représentation des structures verbo-nominales et du figement verbal dans deux formalismes de grammaire fonctionnelle* (mai 2002)
- n° 9 Valeriano BELLOSTA von COLBE & Jacques FRANÇOIS, *Windowing of attention and the core-periphery boundary* (septembre 2002)
- n° 10/1 Maxi KRAUSE, *AUS, élément prépositionnel, élément constitutif de particules diverses et particule verbale* (septembre 2002)
- n° 10/2 Maxi KRAUSE, *VON, préposition et élément constitutif d'adverbes et de tournures adverbiales* (octobre 2002)
- n° 10/3 Maxi KRAUSE, *BEI, préposition, élément constitutif de particules diverses et particule verbale* (novembre 2002)
- n° 10/4 Maxi KRAUSE, *MIT, préposition, élément constitutif de particules diverses et particule adverbiale* (novembre 2002)
- n° 11 Richard RENAULT, *Morphosyntaxe des participes finnois* (décembre 2002)
- n° 12 Dominique LEGALLOIS (coordination), *Unité(s) du texte* [Journée Scientifique du CRISCO, 6 décembre 2002] (juin 2003)
- n°13 Jacques FRANÇOIS (coordination), *Une grammaire de l'interface entre syntaxe, sémantique et pragmatique : la « Role and Reference Grammar » autour de R.D. van Valin Jr.* [Journée Scientifique du CRISCO, 14 mars 2003] (septembre 2003)