

Utilisation d'un corpus catégorisé pour l'étude et la représentation de la synonymie en contexte.

Jean-Luc MANGUIN
CRISCO, CNRS UMR 6170
Université de Caen, 14032 CAEN Cedex

Introduction.

Le but de cet article est de montrer qu'un corpus catégorisé comme Frantext peut constituer l'un des supports de l'étude d'une relation paradigmatique avec des visées pédagogiques ou comparatives. Le principe de la méthode est de mettre à l'épreuve les données relationnelles fournies par l'autre fondement de la méthode (un dictionnaire spécialisé) en les confrontant avec les cooccurrences repérées dans le corpus grâce à la catégorisation des unités lexicales et selon un schéma syntaxique précis. Il est évident que cette catégorisation est insuffisante pour détecter certaines fonctions syntaxiques, mais elle permet dans notre cas de capter la fonction d'adjectif épithète qui nous servira d'exemple. Nous montrerons que notre procédé aboutit à des représentations graphiques d'un grand intérêt pédagogique, ainsi qu'à des mesures quantitatives caractérisant certains emplois de l'unité étudiée. L'exemple choisi est l'adjectif *curieux*, qui est suffisamment courant pour fournir un nombre de données autorisant certains tests statistiques, et qui possède un intérêt linguistique dans le changement de sens induit par son changement de position dans certains contextes nominaux ; par exemple, dans le cas de la synonymie que nous avons choisi d'étudier, *un curieux regard* est probablement *un regard intéressant*, tandis qu'*un regard curieux* sera plutôt *un regard intéressé*.

1. Les données paradigmatiques

Celles-ci sont issues d'un dictionnaire de synonymes construit au laboratoire CRISCO à partir des données provenant de sept dictionnaires de référence, et déjà mentionné dans d'autres publications (le modèle est décrit dans Ploux et Victorri, 1998) ; l'exploitation des données présentes dans ce dictionnaire se fait grâce aux outils de la théorie des graphes : nous formalisons le dictionnaire comme un graphe dont les sommets sont constitués par les mots-vedettes du dictionnaire, et dont les arêtes sont les relations existant entre ces mots (pour la terminologie relative aux graphes, voir Berge, 1958). La relation ainsi définie est réflexive, et symétrique ; ce formalisme a été utilisé en premier par Kahlmann, 1975. Pour étudier une unité, on se limite au sous-graphe dont l'ensemble des sommets est composé de l'unité étudiée et de ses synonymes, et dont l'ensemble des arêtes est constitué par toutes les relations existant entre ces sommets (la variabilité du sous-graphe se rencontre dans Manguin et Victorri, 1999). Pour l'adjectif *curieux*, le sous-graphe ainsi défini possède l'allure suivante :

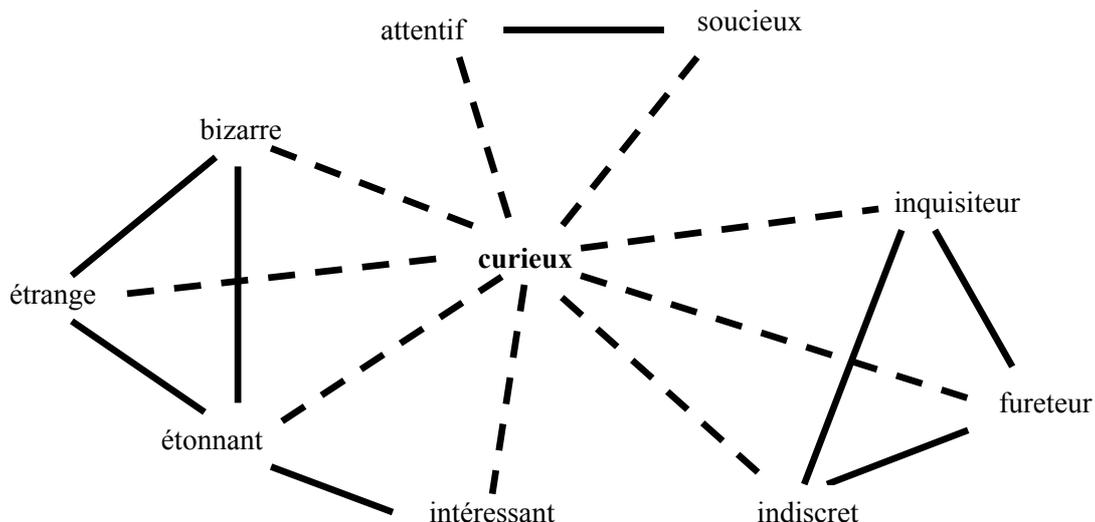


Figure 1 : Structure simplifiée du sous-graphe de synonymie de *curieux*

Dans la réalité de notre dictionnaire, *curieux* possède 34 synonymes adjectivaux, que l'on peut, si l'on exclut du graphe les arêtes passant par le mot-vedette étudié, séparer en trois composantes connexes, que nous donnons ci-après :

1. amusant, attachant, bizarre, drolatique, drôle, déconcertant, extraordinaire, extravagant, incompréhensible, inconcevable, incroyable, inouï, intéressant, original, paradoxal, piquant, pittoresque, plaisant, rare, singulier, surprenant, unique, étonnant, étrange
2. chercheur, fureteur, indiscret, inquisiteur, investigateur
3. anxieux, attentif, avide, intéressé, soucieux

Nous voyons que les composantes 2 et 3 correspondent au sens *intéressé* de notre adjectif, tandis que la composante 1 reflète le sens *intéressant*.

Pour construire la représentation graphique à laquelle nous voulons aboutir, nous devons établir une relation valuée entre les synonymes ; Brodda et Karlgren, 1969, ont déjà proposé une méthode répondant à cette préoccupation, mais elle est relativement délicate à mettre en œuvre car elle doit s'appliquer à l'ensemble du graphe formé par le dictionnaire entier. Notre démarche consiste à partir de la matrice d'adjacence du sous-graphe, et pour chaque paire de sommets A et B, nous définissons une similitude entre ces deux sommets par l'indice de Jaccard S (voir Legendre et Legendre, 1998) :

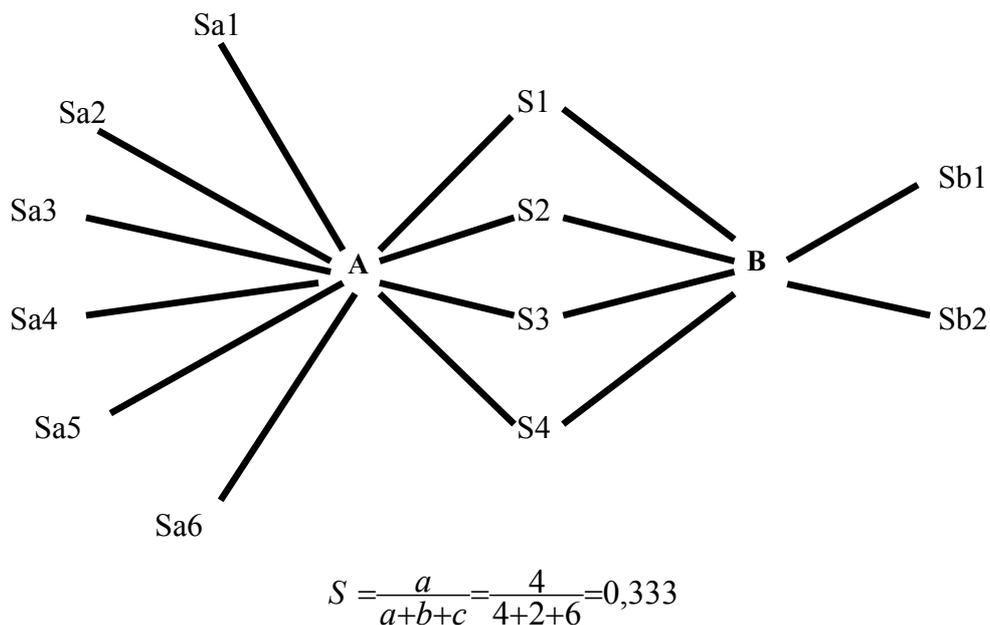


Figure 2 : exemple de graphe illustrant le calcul de la similitude entre sommets

Dans la formule de calcul, a est le nombre de sommets en relation avec A et avec B, b (resp. c) le nombre de sommets en relation avec A mais pas avec B (resp. avec B mais pas avec A). En effectuant ce calcul pour chaque paire de sommets, nous obtenons une matrice de proximités, à partir de laquelle, par échelonnement multidimensionnel (non-metric multidimensional scaling), nous pouvons construire une représentation bidimensionnelle dans laquelle les synonymes s'agencent en fonction des valeurs des proximités (les calculs sont effectués par le logiciel UCINET 6.0, selon une méthode proposée par Kruskal et Wish, 1978). La représentation de *curieux* et de ses 34 synonymes, que nous donnons en annexe (figure 3), montre ainsi un espace clairement séparé en trois zones :

- l'une contenant les synonymes apparentés à *intéressant*, comme *étonnant*, *étrange*, *paradoxal*, etc.
- une deuxième qui inclut les synonymes comme *intéressé*, *attentif* ou encore *soucieux*.
- une troisième où ce second sens est parfois marqué par une appréciation négative, avec des termes comme *indiscret*, *fureteur*, ou *investigateur*.

Il est tout à fait logique que ces trois zones correspondent aux trois composantes connexes signalées plus haut, puisque l'absence de liaison entre deux sommets qui appartiennent à deux composantes connexes différentes induit une similitude nulle, et par conséquent une distance maximale, entre ces deux sommets. En outre, les deux dernières zones se situent dans une moitié de la représentation, opposées à la première zone, et contiennent des termes synonymes de *curieux* quand celui-ci est en position postposée et s'applique à des substantifs « potentiellement actifs » ; autrement dit, on retrouve sur cette représentation la coupure principale entre les deux sens *intéressé* (dans la partie droite) et *intéressant* (dans la partie gauche).

2. Les données syntagmatiques

Pour étudier l'adjectif *curieux* en épithète et en contexte nominal, nous faisons appel à un corpus qui nous permet de repérer simplement les groupes Nom + Adjectif ou Adjectif + Nom ; nous avons utilisé la base Frantext catégorisée, en excluant le genre poétique, entre 1830 et 1999, ce qui correspond à un corpus de 1711 textes et 120 millions de mots.

Le recueil des données se déroule en deux étapes. Tout d'abord, nous recherchons dans la base Frantext catégorisée les occurrences de *curieux* avec un substantif quelconque, au moyen des requêtes suivantes :

$\&e(g=A\ c=\&mcurieux)\ \&e(g=S)$ (antéposition)

$\&e(g=S)\ \&?(\&e(g=Adv)\ \&e(g=A\ c=\&mcurieux))$ (postposition)

On peut remarquer au passage qu'il est nécessaire d'inclure dans la requête des occurrences de *curieux* postposé un adverbe éventuel entre le substantif et l'adjectif. Après rapatriement des 3500 résultats donnés par Frantext, nous effectuons un traitement local par des programmes élaborés au laboratoire par nos soins, et qui permettent de lemmatiser les substantifs rencontrés, afin d'établir la liste des substantifs les plus fréquents employés avec *curieux*, et que nous donnons dans le tableau suivant :

| nom | nb. d'occurrences | % sur 3531 occurrences |
|-----------|-------------------|------------------------|
| chose | 262 | 7,42 % |
| regard | 100 | 2,83 % |
| fait | 66 | 1,87 % |
| phénomène | 65 | 1,84 % |
| spectacle | 54 | 1,53 % |
| détail | 51 | 1,44 % |
| œil | 45 | 1,27 % |
| esprit | 42 | 1,19 % |
| mélange | 39 | 1,10 % |
| histoire | 35 | 0,99 % |
| bête | 30 | 0,85 % |
| cas | 29 | 0,82 % |

Tableau 1 : les 12 premiers substantifs associés préférentiellement à *curieux*

Pour notre étude, nous avons en définitive retenu 25 substantifs, qui sont : *chose, regard, fait, phénomène, spectacle, détail, œil, esprit, mélange, histoire, bête, cas, exemple, livre, type, expérience, personnage, étude, objet, tête, travail, figure, conversation, spécimen* et *page*. On peut noter au passage que cette liste contient plusieurs noms comme *regard, œil, esprit* ou *bête* avec lesquels *curieux* pourra changer de sens suivant sa position, comme signalé en introduction.

Munis de cette liste de substantifs, et de la liste des synonymes, nous créons tout d'abord deux listes (« adj » et « nom »), puis nous interrogeons à nouveau la base Frantext, afin d'obtenir cette fois les cooccurrences nécessaires à notre méthode. Cette fois les requêtes sont de la forme :

$\&e(g=A\ c=\&ladj)\ \&e(g=S\ c=\&lnom)$ (antéposition)

$\&e(g=S\ c=\&lnom)\ \&?(\&e(g=Adv)\ \&e(g=A\ c=\&ladj))$ (postposition)

Nous recueillons ainsi un total de 6547 cooccurrences, dont 1095 concernent *curieux*, et dont le dépouillement automatique selon la même méthode que précédemment nous permet d'aboutir à un tableau (synonymes x substantifs) où chaque case contient le nombre de cooccurrences rencontrées pour la paire envisagée. Nous avons dans notre exemple un tableau à 35 lignes (*curieux* et ses synonymes) et 25 colonnes (substantifs), dont de nombreuses cases sont vides, puisque sur les 1750 paires Nom+Adjectif possibles, seules 414 sont effectivement attestées dans notre corpus, ce qui reflète ce que Coseriu nomme les « solidarités lexicales » de type « sélection » (Coseriu, 2001, pp. 329-330). Ce tableau peut s'assimiler très facilement à la matrice d'adjacence d'un graphe valué, si l'on considère que ce graphe représente la relation de cooccurrence, et que chaque arête porte comme valeur le nombre de cooccurrences pour la paire considérée.

3. Les similitudes d'emploi

De la même manière que nous avons défini une « similitude paradigmatique », nous pouvons à présent définir une « similitude syntagmatique » ou « similitude d'emploi » entre les synonymes à partir des données du tableau des cooccurrences. Toutefois, comme il s'agit de comparer les préférences d'emploi de chaque adjectif avec chaque substantif de la liste, il faut pondérer les données pour ne pas tenir compte de la fréquence absolue de chacun des adjectifs ; nous ramenons ainsi chaque valeur du tableau à un pourcentage calculé en divisant cette valeur par le nombre total d'occurrences de l'adjectif. Nous calculons ensuite la similitude d'emploi en utilisant un indice adapté aux données numériques (en l'occurrence l'indice de Steinhaus, cf. Legendre et Legendre, 1998). Cet indice est basé sur le même principe que celui de Jaccard, c'est-à-dire qu'il divise la partie commune (ici, pour les deux synonymes dont on calcule l'indice, le nombre minimum des cooccurrences observées avec chaque substantif) par la partie totale (le nombre de cooccurrences relevées avec chaque substantif et l'un ou l'autre des deux synonymes), selon le schéma explicatif ci-dessous :

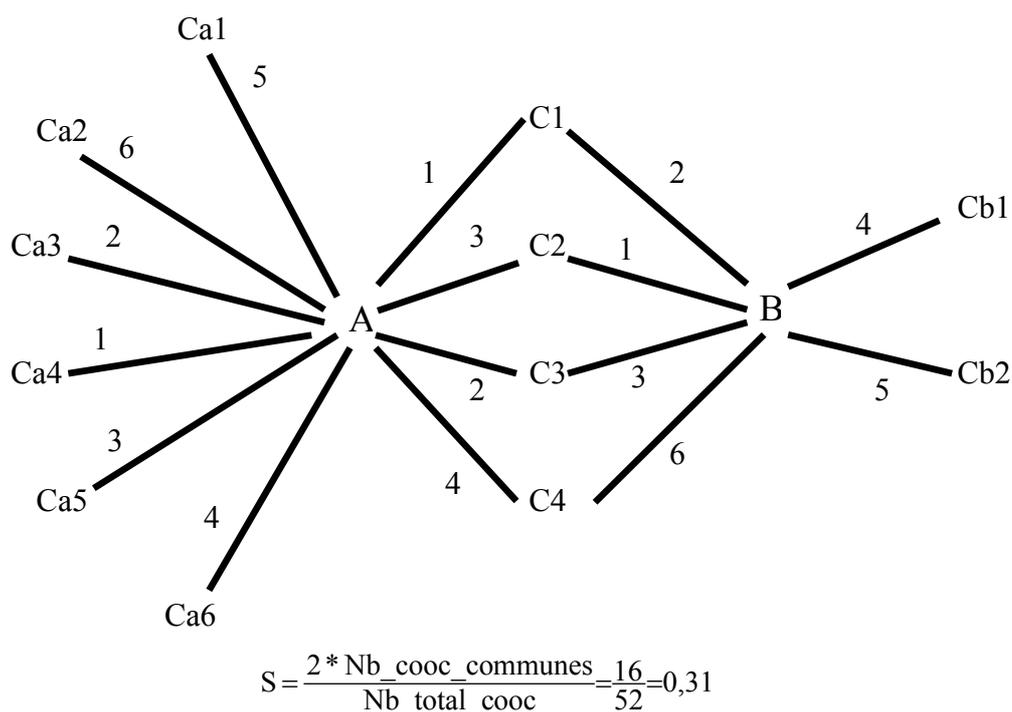


Figure 4 : exemple de graphe illustrant le calcul de la similitude d'emploi

Pour l'exemple présenté sur la figure 4, le calcul de l'indice se fait ainsi :

- au numérateur, on cherche la valeur minimum des deux arêtes partant d'un contexte commun (par exemple pour C4, le minimum est 4), puis on fait le total de ces minima.
- au dénominateur, on totalise les valeurs des arêtes partant de A ou de B.

A l'issue de cette étape, nous sommes en mesure d'appliquer à notre tableau la même transformation non-linéaire, qui nous conduit à une représentation plane situant les synonymes les uns par rapport aux autres (voir figure 5 en annexe) ; ce qui importe ici, c'est la confrontation avec les données paradigmatiques. Nous observons en effet que la dissociation entre le domaine *intéressant* et le domaine *intéressé* existe toujours, et qu'aucun terme n'a changé de zone. Par contre, la dichotomie présente au sein de la zone *intéressé* entre les termes dépréciatifs (comme *furtif*) et les termes neutres (comme *intéressé*) n'est plus présente. Cette différence est probablement due au fait que les données syntagmatiques ne permettent sans doute pas, pour des raisons qu'il importe de préciser, d'accéder à l'essentiel de l'information paradigmatique ; cette conclusion est fondée sur le fait que les regroupements obtenus par classification hiérarchique ascendante des similitudes d'emploi présentent parfois des anomalies sémantiques, si on les compare aux groupes que l'on obtient à partir des similitudes paradigmatiques. Nous pouvons en effet effectuer une classification hiérarchique à partir des tableaux de similitudes, comme nous l'avons déjà présenté par ailleurs (voir Manguin, 2004). Il est vrai que dans le cas

présent, la seule information paradigmatique incluse dans les données est la relation de synonymie entre *curieux* et ses synonymes (puisque nous avons conservé la liste de ceux-ci), et que les relations entre ces synonymes ont disparu. Des groupes comme (*inquisiteur, anxieux*) ou (*avide, investigateur*), issus des similitudes d'emploi, nous paraissent en effet moins cohérents que (*avide, intéressé*) et (*inquisiteur, fureteur*) qui eux proviennent des similitudes paradigmatiques. La solution à ces différences consiste à pondérer les données syntagmatiques par les données paradigmatiques, en faisant simplement case à case la moyenne arithmétique des deux tableaux de similitudes ; on obtient alors une représentation plus conforme à la réalité des données de chaque point de vue, qui ressemble beaucoup à celle réalisée avec les seules similitudes d'emploi, mais cette fois les classifications hiérarchiques construites avec ces similitudes moyennes ne font pas apparaître de regroupements aberrants.

D'un autre côté, les valeurs des similitudes d'emploi peuvent révéler certains choix lexicographiques contestables et par là même invalider un lien synonymique avec le mot étudié si ces valeurs s'avèrent trop faibles. Par exemple, nous avons éliminé de notre liste de synonymes des adjectifs qui avaient des similitudes d'emploi extrêmement faibles avec *curieux*, comme *sensationnel* (sim = 0,08) et *désireux* (sim = 0,06). Le premier n'est mentionné que par un seul de nos dictionnaires sources (voir la liste de nos dictionnaires dans la bibliographie), tandis que le second correspond à un sens qualifié de « vieux » de notre adjectif *curieux*. En outre, la répartition en deux groupes de synonymes au lieu de trois laisse penser qu'une connexité entre tous les termes plus ou moins synonymes d'*intéressé* refléterait peut-être mieux la réalité du paradigme étudié.

4. La représentation du contexte

Sur la représentation graphique précédemment construite, et où figurent les synonymes, il est facile de placer les substantifs dont nous avons étudié les cooccurrences. L'emplacement d'un substantif sera le barycentre des points correspondant aux synonymes qui se rencontrent avec ce substantif, affectés chacun de leur nombre de cooccurrences respectif. Il nous faut cependant effectuer une pondération en divisant, pour chaque substantif S, le nombre de ses cooccurrences avec un adjectif par le nombre total d'occurrences de ce substantif avec tous les adjectifs de la liste.

Le placement de ces substantifs fait parfaitement ressortir leurs « classes » (voir figure 6 en annexe) ; en effet, comme nous l'avons déjà dit, nous avons une représentation où figurent d'un côté les synonymes autour d'*intéressant*, et de l'autre ceux proches d'*intéressé* ; les substantifs se placent eux aussi dans cet espace en deux zones bien séparées :

- côté *intéressé* : *œil, regard, étude, bête, esprit, tête*.
- côté *intéressant* : *cas, fait, exemple, expérience, histoire, conversation, détail* (entre autres).

Pour ne pas surcharger la représentation, nous n'avons pas placé toutes les étiquettes des points correspondant aux substantifs ; par ailleurs, le « groupe des cinq » ne se situe pas exactement dans la zone d'*intéressé*, mais plutôt à mi-chemin d'*intéressé* et d'*intéressant* ; cette position intermédiaire s'interprète comme reflétant le possible changement de sens de *curieux* en position postposée avec ces substantifs (voir le paragraphe 5 qui explique cette conclusion).

Au point de vue lexicologique, il est intéressant d'analyser en détail la position de chacun des termes ; dans le cas d'*étude*, par exemple, le changement de position de *curieux* peut induire non seulement un changement de sens de l'adjectif, mais si c'est le cas, il provoque en plus un changement de sens radical du substantif ; en effet, avec *curieux* antéposé, *étude* correspond au produit d'une action intellectuelle, tandis qu'avec *curieux* postposé, *étude* peut désigner l'action intellectuelle elle-même. Pour les autres substantifs de ce groupe, le changement de position de *curieux* n'entraîne qu'une différence de nuance pour le sens du substantif.

5. Remarques sur la position de l'adjectif *curieux* en épithète

Comme nous l'avons déjà dit, l'adjectif *curieux* possède, dans certains contextes, la particularité de changer parfois de sens en changeant de position par rapport à son substantif régissant. En séparant les données des cooccurrences de *curieux* en deux groupes suivant sa position, nous pouvons calculer comme précédemment, pour chaque synonyme, les deux indices de similitude d'emploi entre ce synonyme et *curieux* en position anté- ou postposée. La comparaison de ces deux séries d'indices montre principalement que la similitude d'emploi avec *curieux* antéposé est un facteur qui différencie les deux groupes de synonymes présents sur la représentation ; de plus, cette différenciation correspond exactement avec la coupure pseudo-homonymique entre les deux sens de *curieux*, comme le montre le tableau ci-après, où les synonymes sont classés par valeur de similitude (avec *curieux* antéposé) croissante :

| synonyme | similitude avec <i>curieux</i> antéposé | similitude avec <i>curieux</i> postposé |
|---------------|--|--|
| inquisiteur | 0,020 | 0,175 |
| chercheur | 0,020 | 0,128 |
| investigateur | 0,037 | 0,222 |
| fureteur | 0,037 | 0,215 |
| avide | 0,053 | 0,278 |
| soucieux | 0,076 | 0,253 |
| intéressé | 0,086 | 0,287 |
| anxieux | 0,087 | 0,255 |
| attentif | 0,149 | 0,293 |
| indiscret | 0,155 | 0,279 |
| inconcevable | 0,248 | 0,394 |
| drôle | 0,263 | 0,557 |
| incroyable | 0,288 | 0,476 |

Tableau 3 : similitudes d'emploi avec *curieux* antéposé ou postposé

La coupure entre *indiscret* et *inconcevable* étant fortement significative, il apparaît clairement que les dix premiers adjectifs de ce tableau ne correspondent jamais au sens que recouvre *curieux* antéposé ; comme nous le verrons dans la conclusion, ce fait peut constituer un préambule à un traitement automatique de la désambiguïsation de *curieux*. Mais d'autre part, il faut aussi tenir compte du fait que ces dix adjectifs ont également une similitude assez faible avec *curieux* postposé, et que seuls *attachant* (0,098), *drolatique* (0,225) et *pittoresque* (0,314) ont des valeurs comparables. Cependant, cette apparente anomalie s'explique très bien si l'on se rappelle que *curieux* postposé porte l'un ou l'autre des deux sens possibles, et que parmi les 25 substantifs retenus, 6 seulement sont susceptibles de donner lieu à un changement de sens de *curieux* quand il est postposé ; en termes de cooccurrences relevées de *curieux*, cela ne représente que 24 %, ou si l'on est plus optimiste, 29 % de celles où *curieux* est en position postposée.

Conclusion et perspectives

Nous avons montré qu'il est possible, en combinant un dictionnaire et un corpus, autrement dit une ressource paradigmatique et une ressource syntagmatique, d'établir une relation valuée entre tous les synonymes d'un même mot-vedette ; en outre, nous avons également montré que cette relation peut aboutir de manière relativement simple à une représentation à but pédagogique. Cette représentation peut de plus s'enrichir de données contextuelles qui contribuent à en améliorer la lisibilité et l'efficacité. Il est également important de constater que nos résultats peuvent constituer un préliminaire à une désambiguïsation automatique du sens de *curieux* en contexte ; comme nous l'avons dit plus haut, les adjectifs situés dans la zone de droite (proches d'*intéressé*) sont les plus éloignés du sens que revêt *curieux* lorsqu'il est antéposé. Le placement d'un substantif dans l'espace de la représentation va permettre de déduire le sens pris par l'adjectif *curieux* ; en effet, si le substantif se place au milieu de la zone correspondant à *intéressant*, alors il n'y a pas d'ambiguïté dans le syntagme Nom + *curieux*. C'est le cas des substantifs comme *objet*, *figure*, *type*, *détail*, etc. En revanche, pour les substantifs comme *esprit*, *tête*, *étude*, *bête*, *œil* et *regard*, leur placement s'interprète ainsi : si ces substantifs sont à mi-chemin des deux zones, c'est parce qu'ils possèdent des solidarités lexicales avec les adjectifs de chaque zone, en particulier celle de la zone *intéressé* ; or ces derniers adjectifs étant très peu similaires à *curieux* antéposé, cela veut dire que les sens qu'ils recouvrent ne peuvent être pris que par *curieux* postposé. Par conséquent, c'est l'emploi de *curieux* postposé avec ces six substantifs qui est seul susceptible d'activer ces sens de l'adjectif étudié ; mais cette activation n'est pas systématique, comme le montre l'exemple suivant où il semble que l'auteur ait cultivé à dessein l'ambiguïté du syntagme *regard curieux* :

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 [R. Martin du Gard, *Les Thibault. Le pénitencier*, p. 792]

Dans cet exemple (analysé dans François, Victorri et Manguin, 2003), l'ambiguïté vient de ce que le co-texte qui précède (*son ami lui devenait étranger*) guide le lecteur vers une interprétation de *regard curieux* en *regard étrange*, tandis que le co-texte qui suit (*un peu moqueur, dont Daniel l'enveloppa*) donne plutôt à *regard* un caractère agentif, et donc une interprétation du syntagme en *regard attentif* ou même *regard inquisiteur*.

Le syntagme *bête curieuse*, habituellement considéré comme figé, illustre le fait que *curieux* ne prend pas toujours le sens *intéressé* quand il est en position postposée ; en effet, dans notre corpus d'étude, nous n'avons pas trouvé d'exemple où la *bête* en question pouvait être considérée comme agent, et les dictionnaires généraux consultés sur ce sujet confirment ce point de vue. On peut néanmoins citer l'article « curieux » du Dictionnaire Historique de la Langue Française : « 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) » (Rey et al., 1992). La position intermédiaire est donc le reflet d'une ambiguïté possible du syntagme Nom + *curieux*, quand le Nom en question est l'un des six que nous avons mentionnés. Un traitement automatique pourrait ainsi repérer les ambiguïtés possibles et confier la résolution de celles-ci à une exploration du texte avoisinant le syntagme repéré comme éventuellement ambigu. Le placement (autrement dit, le calcul des coordonnées) d'un nouveau substantif différent des 25 déjà étudiés se ferait, dans cette méthode, par le relevé des cooccurrences de ce substantif avec les synonymes de *curieux*, puis par calcul du barycentre comme nous l'avons fait pour les substantifs de l'étude.

Dictionnaires sources

BAILLY, René (1946) : Dictionnaire des synonymes. Paris, Larousse.

BÉNAC, Henri (1956) : *Dictionnaire des synonymes*. Paris, Hachette.

BERTAUD DU CHAZAUD, Henri (1971) : *Nouveau dictionnaire des synonymes*. Paris, Robert.

GUIZOT, François (1864) : *Dictionnaire Universel des synonymes de la Langue Française*. Paris, Didier (7ème édition).

LAFAYE, Pierre-Benjamin (1858) : *Dictionnaire des synonymes de la Langue Française*. Paris, Hachette.

Grand Larousse de la Langue Française (1971) : Paris, Larousse.

Le Grand Robert, dictionnaire de la langue française (1985) : sous la dir. d'A. REY. Paris, Robert.

Bibliographie

BERGE, Claude (1958) : *Théorie des graphes et ses applications*. Paris, Dunod.

BORGATTI, S.P., M.G. EVERETT et L.C. FREEMAN (1999) : *UCINET 6.0 version 1.00*. Natick MA, Analytic Technologies.

BRODDA, Benny et Hans KARLGREN (1969) : « Synonyms and synonyms of synonyms », *SMIL*, 5, p. 3-17. Stockholm.

COSERIU, Eugenio (2001) : *L'homme et son langage*, Paris, Peeters.

FRANÇOIS, Jacques, Bernard VICTORRI & Jean-Luc MANGUIN (2003) : « Polysémie adjectivale et synonymie: l'éventail des sens de *curieux* », in Soutet O. (ed.) *La polysémie*, Paris, Presses de l'Université de la Sorbonne.

Frantext, base textuelle catégorisée (1999) : CNRS, ATILF (Analyse et traitement informatique de la langue française), UMR CNRS-Université Nancy2, <http://www.inalf.fr/atilf>.

KAHLMANN, André (1975), *Traitement automatique d'un dictionnaire de synonymes*, Stockholm, Université de Stockholm.

KRUSKAL, Joseph B. et Myron WISH (1978) : *Multidimensional scaling*, Beverly Hills, Sage Publications.

LEGENDRE, Pierre et Louis LEGENDRE (1998) : *Numerical Ecology*, Amsterdam, Elsevier.

MANGUIN, Jean-Luc (2004, à paraître) : « Regroupements de synonymes par indices de similitude : exemple avec l'adjectif *ancien* ». Colloque *Les adjectifs non prédicatifs*, 28-29 novembre 2002, Université de Paris-Nord Villetaneuse. A paraître dans *Les Cahiers de Lexicologie*.

MANGUIN, Jean-Luc & Bernard VICTORRI (1999) : « Représentation géométrique d'un paradigme lexical », *actes de la conférence TALN 1999*, pp. 363-368.

PLOUX, Sabine & Bernard VICTORRI (1998), « Construction d'espaces sémantiques à l'aide de dictionnaires de synonymes », *TAL*, Vol 39/1, pp. 161-182.

REY, Alain (éd. 1992). *Dictionnaire historique de la langue française*. Editions Le Robert.

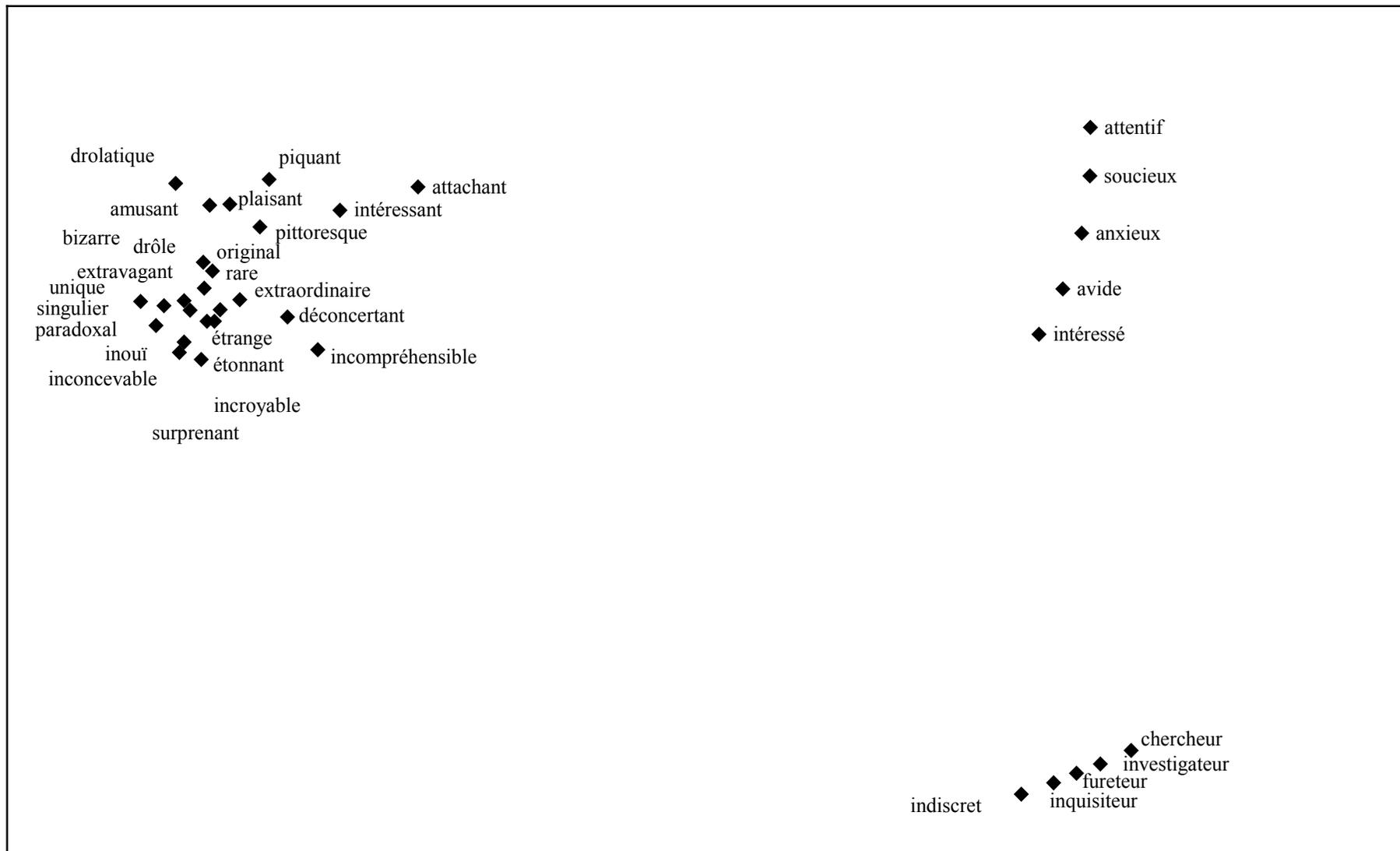


Figure 3 : Représentation de l'espace sémantique de *curieux* par positionnement multidimensionnel, d'après les similitudes paradigmatiques

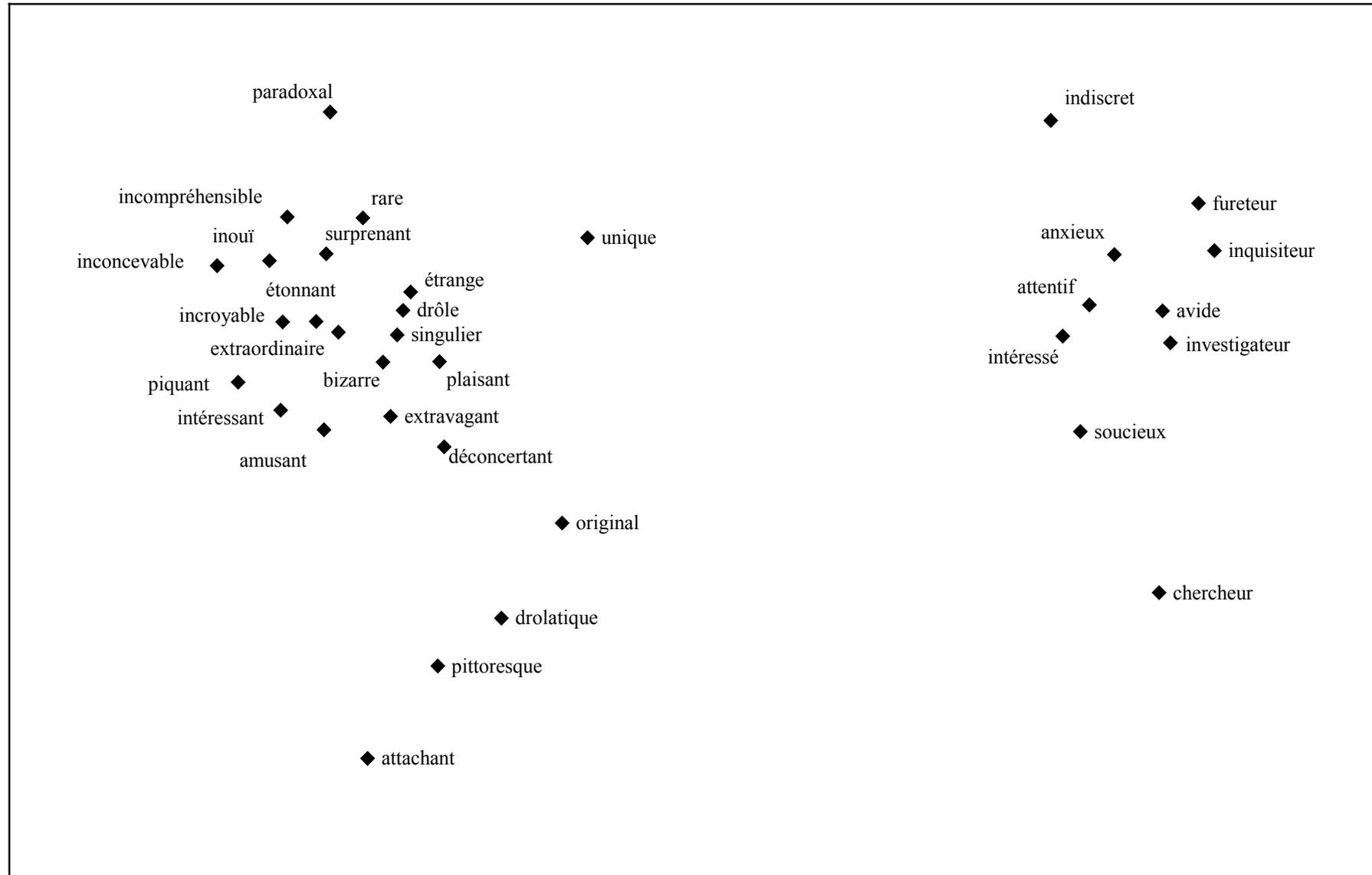


Figure 5 : Représentation de l'espace sémantique de *curieux* par positionnement multidimensionnel, d'après les similitudes d'emploi

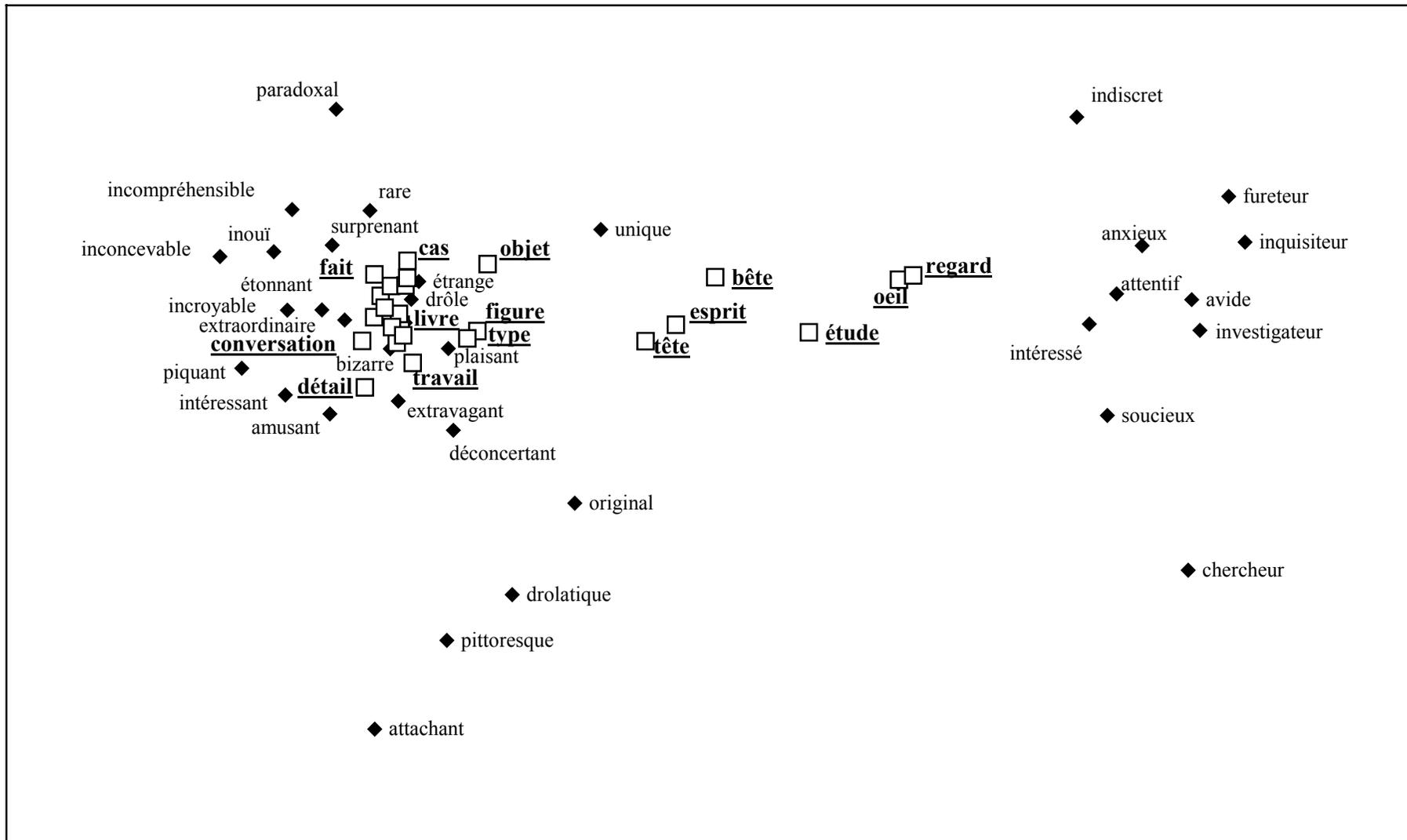


Figure 6 : Représentation de l'espace sémantique de *curieux* et de ses contextes nominaux préférés, d'après les similitudes d'emploi