

# **COMPARAISON DE METHODE: A PROPOS DE L'ARTICLE "APPORT DES MODELES GRAPHIQUES GAUSSIENS EN ANALYSE DE SIMILITUDE"**

**PIERRE VERGES**

CNRS, Aix en Provence, France

**CLAUDE FLAMENT**

Université de Provence, Aix en Provence, France

Modèle gaussien et Analyse de similitude sont mis en perspective dans l'article de Nicolas Roussiau, Saïd Jmel et Joseph Saint-Pierre parce qu'ils donnent lieu tous les deux à une représentation graphique des données. Mais avant d'entrer dans un débat de méthodes il convient de s'interroger sur les données de leur exemple: dans quelle mesure sont-elles analysables sans déformation par de telles méthodes ? En effet, une hypothèse forte du modèle gaussien suppose que les données soient distribuées selon la loi de gauss. Or, fondamentalement les réponses à un questionnaire du type de celui présenté dans cet article, ne sont pas « normales » (au sens de suivre la loi de gauss) et plus généralement les données des questionnaires de représentations sociales ne visent pas à être gaussienne.

On observe ici trois types de distribution. On a d'abord une série d'items pour lesquels les notes extrêmes du type 5, 6 ou 1, 2 sont fortes c'est-à-dire où on trouve plus de 50 % des sujets ayant répondu dans les positions extrêmes de l'échelle. Nous appelons ces items des items en « J » dans la mesure où leur distribution comme le montre la figure n°1 ne répond absolument pas au modèle gaussien.

Une seconde série d'items est à peu près semblable à l'exception près que les valeurs extrêmes, la valeur 6 ou la valeur 1, n'est pas choisie par les sujets. Leurs réponses se concentrent sur les réponses 4-5 ou 2-3. La courbe en J est alors un peu déformée. F.H.Allport montre dans son article sur la courbe en J [Allport, F.H., 1934] que la multiplication des échelons tend à faire disparaître, même dans les situations d'observation de comportements, la forme en J. Si les enquêteurs avaient proposé 3 ou 4 échelons, elle aurait été plus nette. Inversement s'ils avaient choisi une échelle en 9 ou 10

échelons on se rapprocherait artificiellement d'une courbe de Gauss ou d'une courbe ayant un fort plateau horizontal.

Ces deux séries d'items concernent 12 items sur 17. Nous renvoyons ici à notre article dans le Cahier international de psychologie sociale [Vergès, 1995] où on montre que les réponses aux questionnaires de représentations sociales proposant des échelles donnent quatre types de distribution dont la distribution gaussienne n'est qu'une parmi d'autres. A la limite cette distribution de type gaussienne ici centrée autour des réponses 3 et 4 indique un item peu caractéristique de l'ensemble du champ que l'on est en train d'étudier (ici le « Politique ») puisque les répondants se réfugient sur les valeurs moyennes, ne prenant position ni d'un côté, ni de l'autre.

TABLEAU N°1  
Distribution des Items par score sur l'échelle (en nombre n=553)

Item	Moyenne	1	2	3	4	5	6
N	5.23	6	5	9	52	246	35
Q	4.76	3	9	35	164	203	139
H	4.75	1	7	28	198	177	142
K	4.72	3	10	35	181	184	140
I	4.32	2	16	53	269	159	54
L	4.24	12	31	80	180	192	58
C	4.19	31	56	61	136	168	101
D	4.13	16	20	85	218	181	33
O	3.05	71	153	110	134	68	17
G	3.03	77	133	130	129	78	6
A	2.97	52	135	176	162	25	3
J	2.85	40	146	237	119	10	1
E	2.84	62	143	188	145	13	2
F	2.41	71	264	148	62	8	0
B	2.30	136	236	103	44	26	8
M	2.17	168	204	120	43	16	2
P	1.89	336	94	35	45	23	20

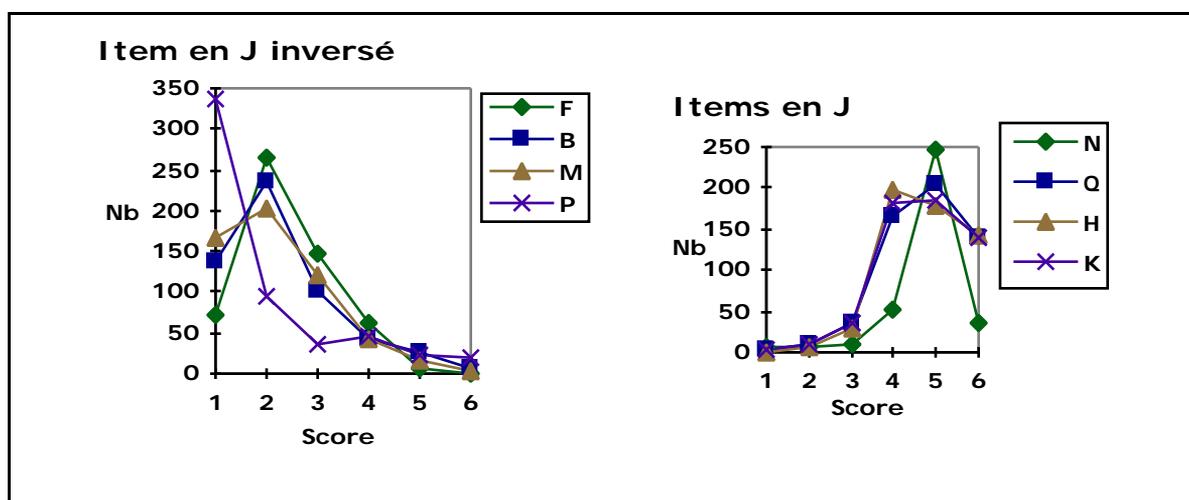


Figure n°1

### 1. UNE ANALYSE DE CORRELATION QUI REPRODUIT LA HIERARCHIE DES MOYENNES

Les données de ce questionnaire ne sont pas systématiquement normales et cette non-normalité va structurer tout indice de corrélation que l'on peut établir entre ces données. En effet, si on ordonne les items selon leur valeur moyenne, on obtient une matrice de similitude (l'indice est ici le  $\tau$  de Kendall) qui n'exprime pratiquement que cette hiérarchie de la moyenne des items. On a par exemple un premier bloc de cinq items très corrélés entre eux et ayant des valeurs moyennes très nettement supérieures à 3 (N, Q, H, K, I). On trouve les six derniers items (en dehors de P) aussi très corrélés entre eux et ayant des moyennes bien inférieures à 3 (A, J, E, F, B, M). La matrice de similitude sauf pour quelques items (item O, P en particulier) ne donnent pas beaucoup plus d'informations que l'ordre donné par la moyenne des réponses.

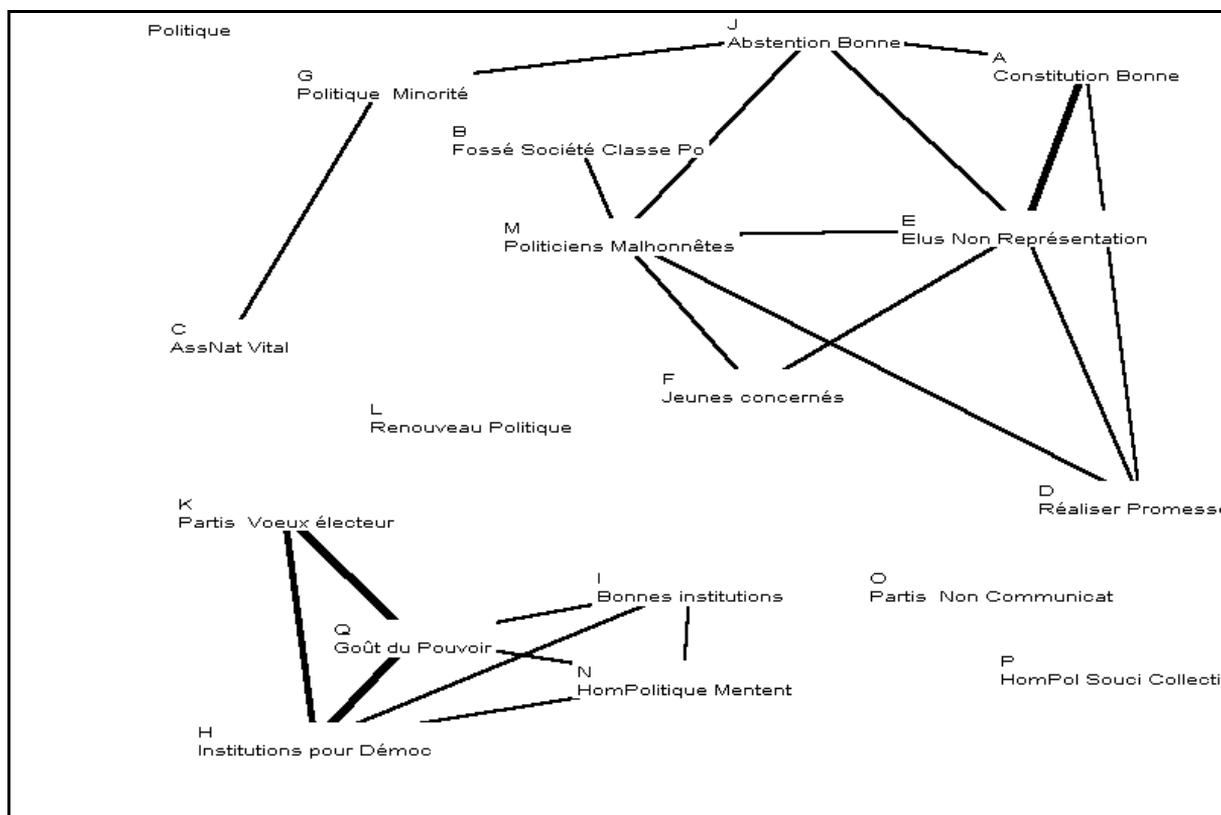


Figure n°2  
 Graphe de similitude: les relations supérieures à 0.15 ( $\tau$  de Kendall)

Le graphe de la figure n°2 est assez fortement déterminé par cette segmentation initiale: les items ayant une courbe en « J » soit centrés sur 5-6, soit centrés sur les valeurs 1-2. Le graphe de similitude est alors composé de deux zones, une zone pour laquelle les sujets ont répondu qu'ils n'étaient pas d'accord avec les phrases, et une zone pour laquelle ils ont répondu massivement qu'ils étaient d'accord avec ces phrases.

Quelle information ces questionnaires nous apportent-ils si on s'en tient uniquement aux chiffres et aux résultats des méthodes de traitement. Ils nous apportent une première information liée au tri à plat, c'est-à-dire aux moyennes obtenues par tous ces items. Aussi la phrase la plus acceptée par les sujets enquêtés concerne négativement la représentation des électeurs (N: *les gens ne se sentent plus correctement représentés par les élus*). Inversement l'item le plus rejeté concerne, en reflet, le comportement des électeurs (P: *s'abstenir c'est ce qu'il y a de mieux à faire un jour d'élection*). mais entre ces deux items il n'y a pas symétrie. Le premier est bien intégré dans une zone dense du graphe. Le second n'est pratiquement pas corrélé avec les autres items. On le retrouve seulement corrélé de façon négative avec l'item C qui porte sur les jeunes et la politique.

Cette première information du tri à plat est un peu modulé par l'analyse de similitude mais pas fondamentalement. On tire au mieux de ces corrélations quelques informations complémentaires.

## 2. COMPARAISON DES METHODES

On est alors en mesure de se demander si l'analyse de similitude et l'analyse gaussienne produisent des graphes différents, si l'analyse gaussienne donne un meilleur résultat. Pour établir cette comparaison nous prenons pour matrice de similitude la matrice des corrélations partielles, base de l'analyse graphique gaussienne. Le graphe n'apparaît pas fondamentalement différent (figure n°3).

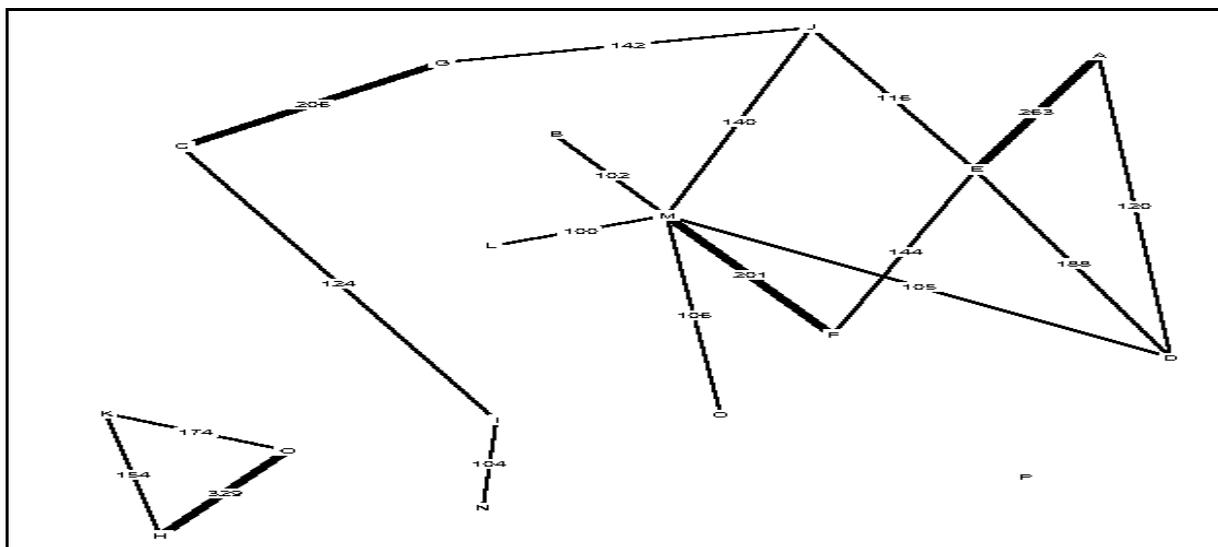


Figure 3

Graphe de similitude: les relations supérieures à 0.10 (Coefficient de Corrélation partielle)

Visuellement il montre quelques différences mais une analyse de corrélation entre ces deux matrices de similitude donne une corrélation de 0.85. On a vérifié que cette concordance n'était pas liée à la différence entre le  $\tau$  de Kendall et le  $\rho$  de Spearman.: la corrélation est entre les matrices de ces deux indices de 0.93. Il y a donc un certain effet

du passage de la corrélation deux à deux à la corrélation partielle mais cet effet est réduit. Il concerne surtout certains items: B (0.72), C (0.82), H (0.82), items qui ont, pour deux d'entre eux, des courbes en « J » très accentuées.

La première explication d'une différence de résultats repose sur l'élimination statistique de ce qui est redevable, dans une corrélation donnée entre deux éléments, aux corrélations de ces deux éléments avec les autres variables. Dans le cas d'une analyse de la relation entre deux variables qui dépendrait fortement d'une troisième variable, cette analyse se justifie pleinement. On cite toujours ce célèbre cas d'école où dans une ville imaginaire les garçons réussissent mieux au Baccalauréat que les filles mais où l'analyse lycée par lycée, montre au sein de chaque lycée la meilleure réussite des filles. Les garçons se trouvent, comme par hasard, dans les lycées ayant la meilleure réussite ! En analyse multivariée le passage de la matrice de corrélation à la matrice de corrélation partielle ne se justifie pas autant car on agrège une multiplicité de relations triangulaires et on en retient une sorte de moyenne. On perd la spécificité de la mise en évidence d'une « variable cachée » qui explique la relation entre deux autres variables ou même l'infirme.

Au lieu de parler d'explication d'une relation par d'autres nous avons toujours mis en avant l'idée d'une similitude « relayée » par un chemin entre deux variables (dans la figure 3 la relation M-C relayée par la chaîne M-J-G-C). Cette idée est à la base de la première procédure de l'analyse de similitude: la 3-analyse. Poussée à l'extrême cette idée aboutit à l'arbre maximum. Celui-ci est constitué de l'ensemble des chemins qui relayent les similitudes entre les variables [Flament, C., 1986]. La corrélation partielle traduit la même notion de manière plus formalisée mais moins intuitive et moins interprétable.

Nous avons pu, grâce à l'amabilité de S. Jmel, travailler sur le détail pas à pas d'une analyse par les modèles graphiques gaussiens d'une matrice de 6 variables. Il est alors possible de comprendre à quel moment analyse de similitude et analyse graphique gaussienne divergent

TABLEAU N°2  
Matrice de corrélations partielles: exemple

	U	V	W	X	Y	Z
U	1	0.37	0.33	0.14	0.03	0.17
V	0.37	1	0.26	0.25	0.01	0.10
W	0.33	0.26	1	0.18	0.01	0.05
X	0.14	0.25	0.18	1	0.07	-0.01
Y	0.03	0.01	0.01	0.07	1	0.61
Z	0.17	0.10	0.05	-0.01	0.61	1

A l'évidence les variables forment deux sous-groupes: U-V-W-X et Y-Z, la seule liaison significative entre ces deux sous-groupes étant U-Z 0.17.

Le pré-ordre de ces corrélations par ordre croissant est le suivant: X-Z -0.01, V-Y 0.01, W-Y 0.01, U-Y 0.03, W-Z 0.05, X-Y 0.07, V-Z 0.10, U-X 0.14, U-Z 0.17, W-X 0.18, V-X 0.25, V-W 0.26, U-W 0.33, U-V 0.37, Y-Z 0.61

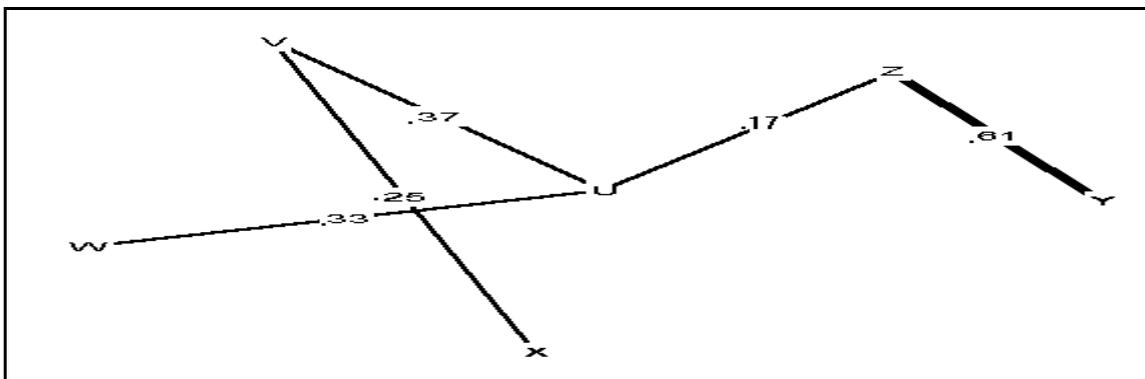


Figure 4  
Arbre maximum de l'exemple.

Si on enlève l'arête la plus faible de l'arbre maximum on obtient les deux composantes connexes vues dans le tableau précédent. On peut au contraire ajouter les arêtes supérieures à un certain seuil. On obtient alors un graphe-seuil comme celui de la figure 5. Ce graphe correspond à l'un des modèles présentés par l'analyse graphique gaussienne.

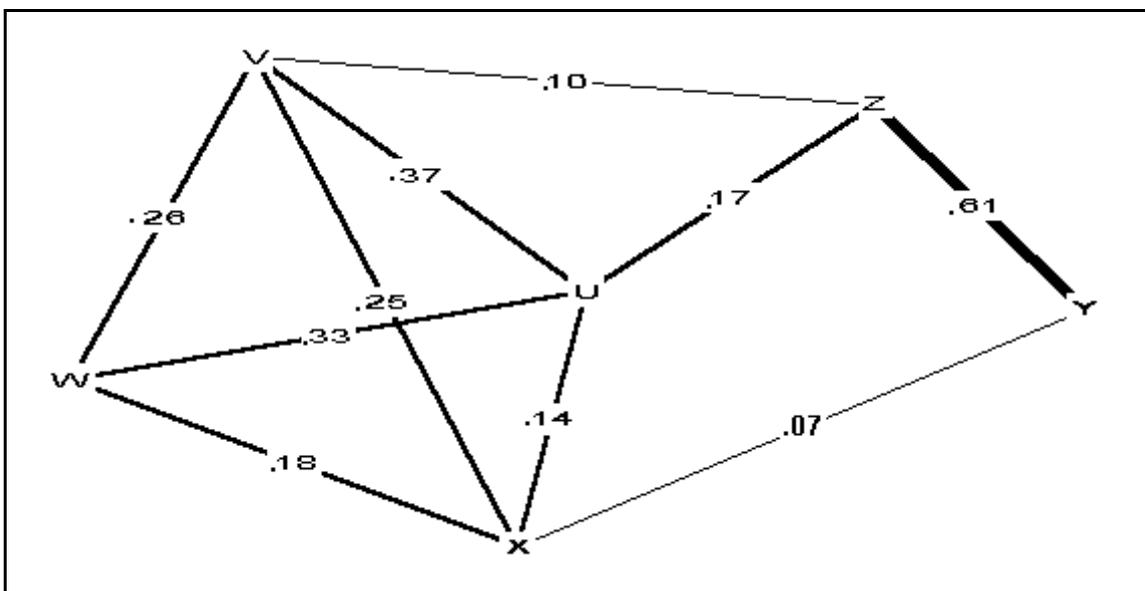


Figure 5  
Graphe de similitude de l'exemple (arêtes supérieures à .05).

L'exploration du pré-ordre de manière décroissante en partant de l'arête la plus forte (Y-Z 0.61) se représente par le filtrant des cliques.

Le filtrant des cliques (limité aux cliques dont la valeur est positive) est alors très éclairant. Il montre comment s'est constitué le graphe au seuil de 0.06. Il montre aussi que les cliques maximales assez significatives (supérieures à .03) donnent presque un autre modèle présenté par l'analyse graphique gaussienne qui a cependant éliminé l'arête W-Z de valeur .05. On peut alors se poser la question de l'élimination de cette arête.

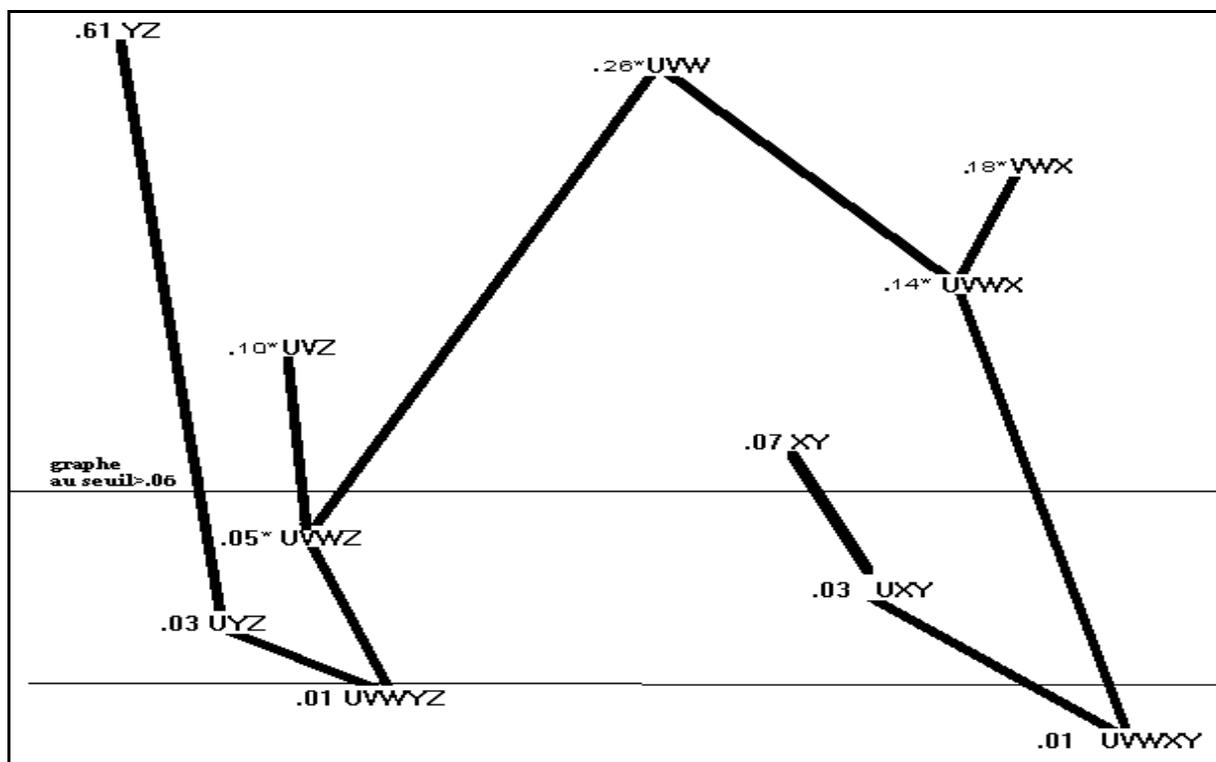


Figure 6  
 Graphe des cliques maximales (Filtrant partiel)

Une procédure de l'analyse graphique gaussienne (Stepwise) fonctionne à partir des valeurs les plus petites. Ou plus exactement elle met à zéro la valeur la plus petite (ici X-Z qui ferme le filtrant sur la clique maximale complète UVWXYZ). Ensuite, elle recalcule la matrice des corrélations partielles et élimine la plus faible arête différente de 0.00. On obtient en éliminant les arêtes les plus faibles (X-Z, V-Y, W-Y, U-Y, W-Z, X-Y) une procédure gaussienne qui suit exactement le pré-ordre de l'analyse de similitude. Mais cette procédure recalcule à chaque pas la matrice de corrélation partielle et ce préordre peut en être partiellement modifié. Avec cette procédure Stepwise on arrive au graphe seuil des arêtes de valeurs supérieure à 0.07 (valeur de la dernière arête enlevé X-Y).

Par contre les autres procédures de l'analyse graphique gaussienne (procédure EH par exemple) donnent 5 modèles dont 2 n'ont pas de correspondance en terme de cliques maximales. Les trois autres correspondent: le premier au graphe de la figure 5 (au seuil de 0.06), les deux autres aux graphes au seuil de 0.05 et 0.03 mais dont on aurait permuté une des arêtes. L'arête enlevée est X-Y (0.07) alors que dans un cas U-Y (0.03) est conservée et dans l'autre W-Z (0.05). La logique de cet entorse au préordre initial revient au recalcul à chaque pas de la matrice de corrélation partielle. La question que l'on doit alors se poser est celle de la correspondance de cette procédure mathématique qui a sa justification au sein de la théorie statistique, avec les hypothèses sur les représentations sociales. A-t-on le droit de modifier un préordre collectif produit par les données recueillies ? La procédure de ce recalcul a-t-elle une interprétation

psycho-sociologique ? Guttman a effectué une telle opération pour obtenir l'axe de son échelle. Dans ce cas l'interprétation est immédiate. Mais ici c'est plus une technique statistique qu'une opération justifiée par la recherche d'un modèle cognitif.

Or entre les résultats que nous obtenons ici et les résultats obtenus dans l'article de Nicolas Roussiau et al. il y a de très grosses différences difficilement acceptables. La quasi totalité des groupes de variables qu'il obtient sont à cheval sur les deux ensembles d'items que nous venons de définir (ACGHN par exemple). A et C par exemple se retrouvent ensemble dans cinq regroupements alors que leur corrélation en termes de Kendall est faible ( $\tau = 0.10$ ; il est de 0.01 en corrélation partielle), que les deux distributions de ces items comme on peut le voir sur le tableau n°1 sont presque inverses. L'explication doit alors être trouvée en analysant dans le détail la procédure suivie par les programmes pour comprendre pourquoi on arrive à de telles différences.

L'intérêt d'une comparaison entre un modèle gaussien et l'analyse des similitudes repose sur la possibilité de comparer deux approches assez fondamentalement différentes. Comme le souligne l'article, le modèle gaussien d'analyse présenté repose sur la possibilité de comparer deux variables, ici des items, non pas entre eux (dans quelle mesure leur croisement indique une corrélation) mais relativement aux corrélations que deux items ont vis-à-vis de tous les autres items (dans quelle mesure ces corrélations ne dépendent pas d'autres corrélations). L'analyse des similitudes, de son côté, ne raisonne que sur les paires, c'est-à-dire sur le tableau de croisement entre deux items et par là même sur la relation qu'exprime ce tableau: cette relation n'est pas obligatoirement une corrélation. Au contraire bien souvent c'est une fréquence (co-occurrence par exemple) ou un indice de distance à l'indépendance statistique ou encore de distance à l'implication. Les références de l'article de Nicolas Roussiau -Abric, 1984, Guimelli, 1989- par exemple n'utilisent pas un indice de corrélation mais un indice de co-occurrence.. La différence entre les résultats du modèle gaussien et de l'analyse des similitudes sont certainement redevables pour une large part, à cette volonté de privilégier d'un côté, un espace global où tous les items sont considérés en même temps, et de l'autre côté, dans l'analyse de similitude, à ne privilégier que les rapports locaux et les valeurs extrêmes (les relations ayant des valeurs de similitude suffisamment fortes) en ne faisant pas dépendre le résultat graphique de relations moins significatives ou n'ayant que peu d'importance. Par exemple cette logique des valeurs extrêmes est poussée à son terme dans la représentation des données par l'arbre maximum: toutes arêtes de cet arbre est première dans son co-cycle<sup>1</sup>. Un autre point où les deux méthodes raisonnent de manière différente est celui de la connexité. Celle-ci n'est pas obligatoire dans l'analyse gaussienne, rien ne l'impose. Au contraire dans l'analyse de similitude le principe même de l'arbre maximum impose la connexité. Ce n'est qu'ensuite qu'on est conduit à explorer le filtrant des cliques qui peut montrer des regroupements partiels de certaines variables, ou des chaînes, ou encore un point d'articulation entre zones fortement connexes. La propriété n'est pas que mathématique. Elle a son correspondant théorique dans la volonté de rechercher une représentation, dans sa complexité et hiérarchisation, et non des morceaux de représentation.

---

<sup>1</sup> Un Co-cycle est composé de toutes les arêtes reliant les sommets des deux sous-arbres qui existeraient si on enlevait une arête de l'arbre.

Il est très intéressant de poursuivre un travail de comparaison entre des méthodes qui, d'un côté, posent le caractère local, extrémal et symétrique de la représentation, et, de l'autre, son caractère beaucoup plus global et dans certains cas causal de la représentation. Ici, il semble qu'il n'est pas possible de dire comment et pourquoi la méthode graphique gaussienne donne des résultats si différents sur la base de matrices comparables.

Par contre, le travail à partir du détail des réponses initiales des sujets que nous avons entrepris au début de cet article, nous paraît absolument essentiel et premier si l'on veut comparer ces types de méthodes et interpréter les apports de l'une et de l'autre. Il nous est alors possible d'interpréter l'apport de l'analyse de similitude. Elle indique dans quelle mesure certains items semblent se situer hors du champ des opinions majoritaires, mais nous ne sommes pas en mesure d'apporter un regard interprétatif sur l'ensemble des éléments rapportés par l'analyse gaussienne, en particulier le rôle de l'item A. Nous pensons qu'avant toute comparaison il est essentiel de réfléchir sur la structure initiale des données qu'on enregistre et sur les courbes de distribution des items.

### **3- ET SI ON ACCEPTAIT DE RETOURNER CERTAINES ECHELLES ?**

Les deux groupes d'items qui ressortent de l'analyse ne sont-ils pas artefactuels ? Ne sont-ils pas une production de la forme positive ou négative des phrases ? La question alors se pose de savoir s'il n'est pas logique de retourner un certain nombre de ces échelles. En effet, on peut prendre l'une des phrases telle que celle où (F). Cette phrase a un score de 2,41, c'est-à-dire que les gens sont assez en désaccord avec celle-ci. Si on retourne l'échelle en considérant la phrase « *les partis n'agissent pas conformément aux vœux de leur électorat* » et si nous faisons l'hypothèse que les sujets se situeraient de manière strictement inverse sur l'échelle proposée, on obtiendra une moyenne de 4,59.

Cette inversion a l'intérêt de gommer le puissant effet de la hiérarchie initiale des scores moyens. Ainsi, on évitera la séparation en deux blocs, le bloc positif et le bloc négatif. On fabrique alors une nouvelle matrice d'intercorrélation avec pour principe que la moyenne de chaque item doit être supérieure à 3 et donc que l'on considère que les réponses sont des réponses d'accords aux différents items. Ainsi on se donne la possibilité d'observer une corrélation positive quel que soit le sens dans lequel la phrase ait été écrite (certaines sont négatives comme H: *les hommes politiques mentent...*, d'autres positives telles que L: *l'Assemblée nationale est un organe vital dans la vie démocratique*) et quel que soit l'opinion des sujets (positive ou négative). On peut cependant se demander si la même phrase écrite au positif et écrite au négatif aurait donné exactement la distribution inverse ? Cette question est centrale alors qu'elle est souvent oubliée. On sait bien qu'il existe une non symétrie de ce type d'inversion car il y a souvent un tiers terme.

Nous nous refusons à retourner une échelle sur le seul fait que cet item est corrélés trop souvent de manière négative aux autres items (cas de D par exemple). On ne cherche pas une matrice maximisant le nombre de coefficients positifs mais effaçant l'opposition entre opinion positive et négative. L'expérience peut cependant être tentée.

Ici nous inversons certaines échelles pour voir si on peut dépasser les premiers résultats de l'analyse de similitude qui ne font (très globalement) que refléter les scores moyens. La matrice de similitude met alors en évidence un regroupement comprenant un

grand nombre d'items, 11 sur 17, qui se trouvent fortement corrélés entre eux. Ces items sont sémantiquement très cohérents: ils expriment une défiance vis-à-vis de la politique. L'item sur lequel l'accord est le plus grand (N: *les gens ne se sentent plus correctement représentés par les élus* se trouve associé à l'item Q: *beaucoup de politiciens deviennent malhonnête* ou K: *c'est le goût du pouvoir qui pousse à se présenter à des élections*. On observe une corrélation tout aussi forte avec l'item M\*<sup>2</sup>: *réaliser les promesses faites aux électeurs n'est pas l'objectif que poursuivent les élus* ou l'item B\*: *le fossé entre la société et la classe politique ne tend pas à se réduire*. A partir de là, la suite ordonnée des items, met en évidence une assez bonne cohérence de la vision de ce champ politique.

L'ordre ainsi construit nous montre une vision relativement négative du champ politique. Et sur ce fond négatif ce sont les hommes politiques qui sont le plus mal vu. Les réalités institutionnelles sont un peu moins touchés telles que (A\*) *la constitution de la 5ème République ne fonctionne pas bien*, (O\*) *la politique ne peut (pas) intéresser qu'une minorité de gens*, et que (I) *les partis politiques sont devenus des organisations où la communication est bloquée*. Ces items ont des scores plus proches de la moyenne (3,00). Donc d'une certaine façon les enquêtés portent plus leur ressentiment sur les hommes politiques que sur un champ politique plus général, un peu plus impersonnel.

Si nous poursuivons plus loin, en construisant et analysant le Filtrant des cliques, la première question est alors celle de la définition des corrélations significatives. En prenant pour base la table statistique standard des écarts de confiance sur les coefficients de corrélation relativement à la valeur zéro, on trouve deux valeurs significatives: 0.115 en acceptant un risque d'erreur de 5 %, et 0.14 en acceptant un risque de 1 %. Ici il est intéressant de comparer ces valeurs à la distribution des valeurs de l'indice de similitude de la matrice de similitude (figure n°4).

Figure n°4: la distribution des valeurs de la matrice de similitude ( de Kendall)

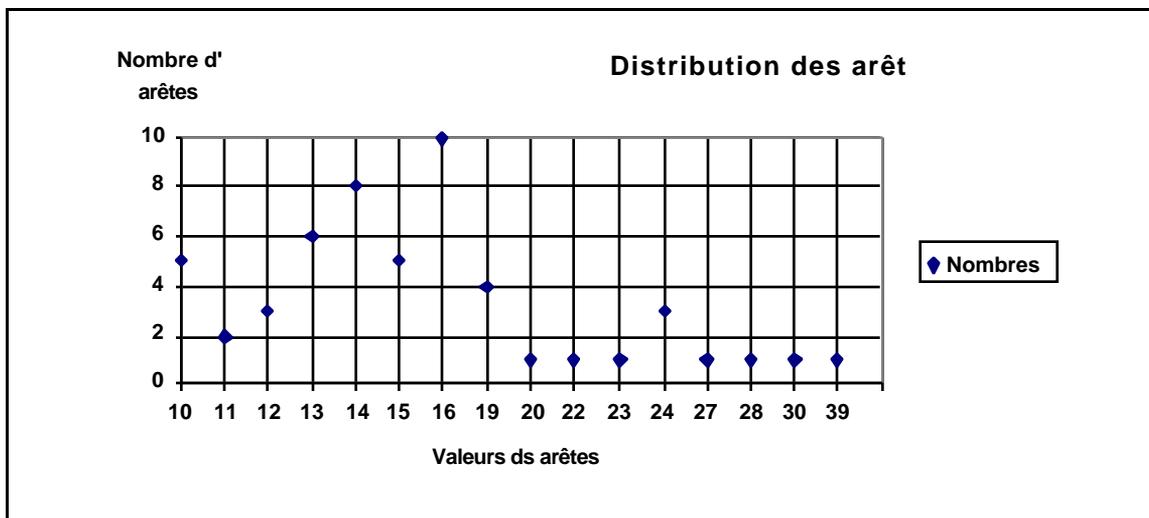


Figure No 4.

<sup>2</sup> L'astérisque indique que l'on a retourné l'échelle des scores de cet item. On indique alors la phrase dans son nouveau sens (négatif si elle était positive et inversement).



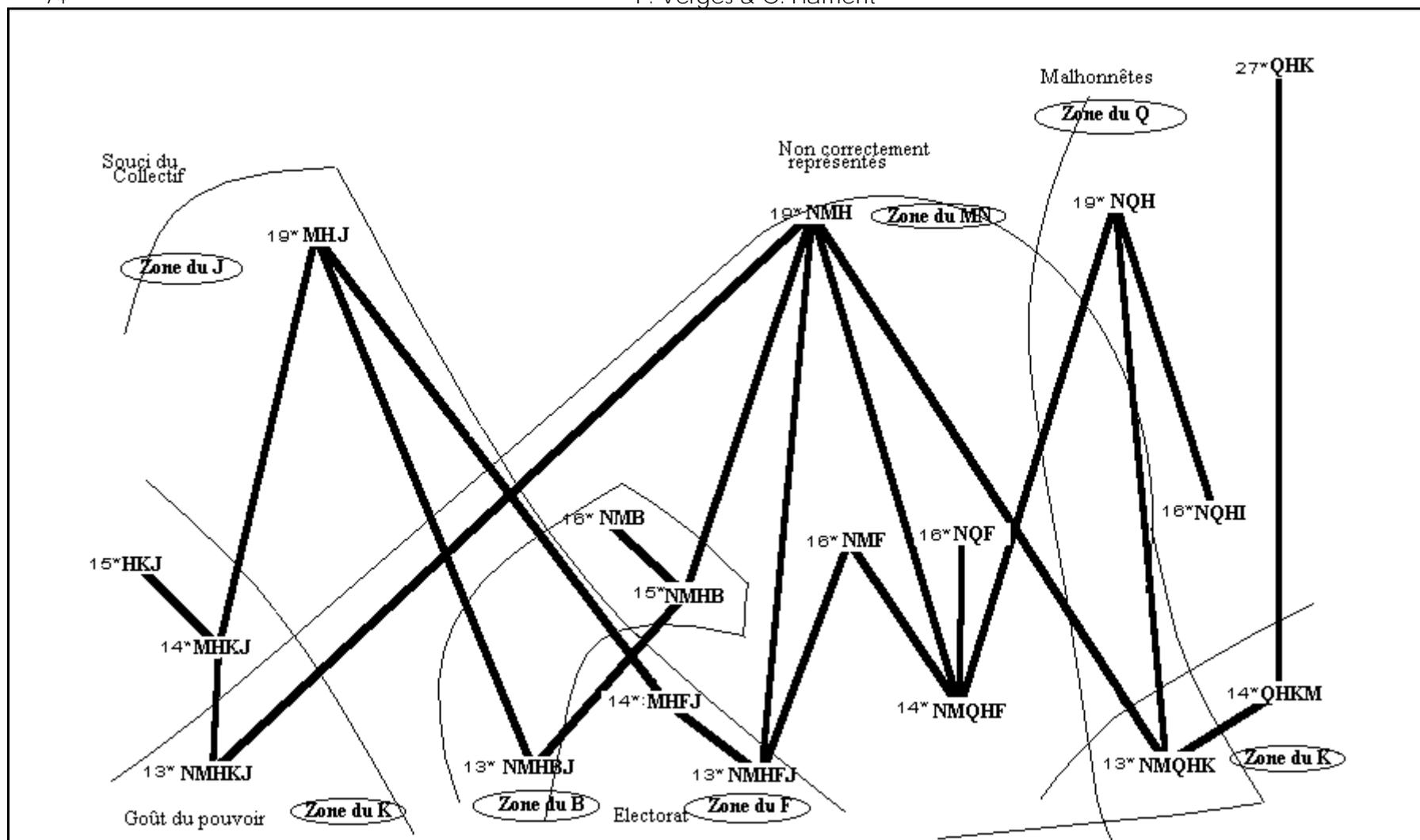


Figure 6  
Les cliques maximales de la zone principale

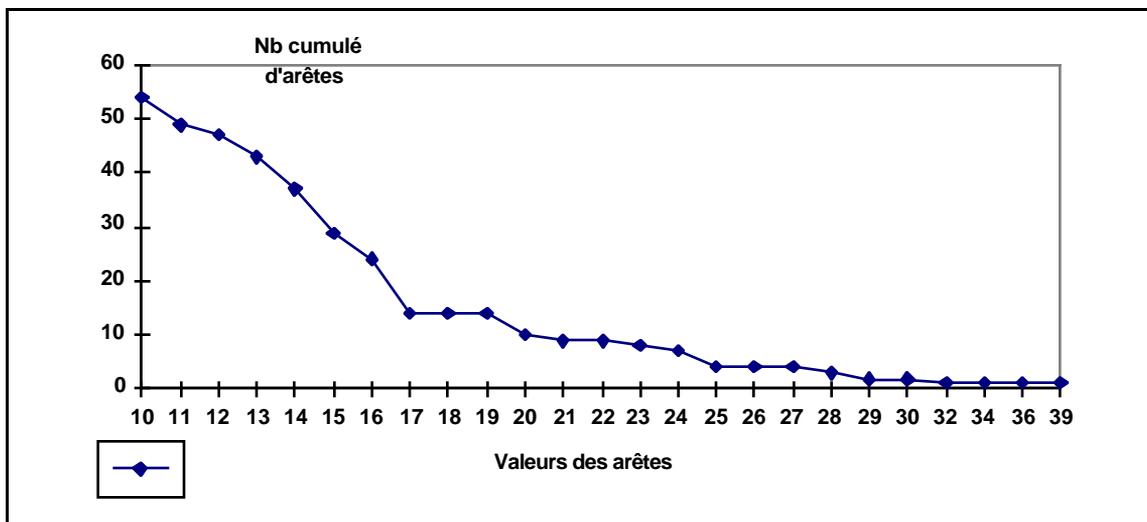


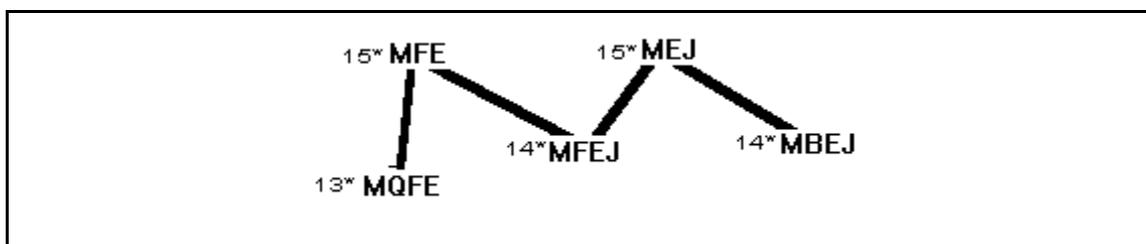
Figure n°5  
 La distribution cumulée des valeurs de la matrice de similitude (  $\tau$  de Kendall)

On observe que jusqu'à 0.20 les différentes valeurs sont représentées d'une manière unique, c'est-à-dire que la valeur 0.39 ne se trouve qu'une fois, de même que la valeur 0.22 ou 0.20. Par contre on assiste à un saut qualitatif important à partir de la valeur 0.19 jusqu'à la valeur 0.14 puisqu'on trouve successivement à 0.17 quatre valeurs, à 0.16 dix valeurs à 0.15 cinq valeurs et à 0.14 six valeurs. La distribution cumulée montre que jusqu'à 0.17 la courbe est presque plate et qu'ensuite de 0.17 à 0.12 la courbe a une forte pente. Ensuite cette courbe a une pente moins accentuée. Il semble alors pertinent de considérer la valeur 0.13 comme un seuil de signification intéressant. La valeur 0.11 ne rajoutant pas beaucoup d'informations par rapport à cette valeur 0.13 puisqu'on ne trouve que quelques arêtes entre ces deux valeurs (Cf. figure n°4).

Le filtrant est bien particulier. Il se développe sous la forme d'un zig-zag de la droite vers la gauche. Un élément est commun à l'ensemble des cliques maximales: c'est l'item H (*les hommes politiques mentent plus souvent qu'ils ne disent la vérité*). En partant de la droite on remarque que l'item Q (*beaucoup de politiciens deviennent malhonnêtes*), est présent seulement dans les cliques les plus à droite, au centre ce sont d'abord les items N et M\* (c'est-à-dire *les gens ne se sentent plus correctement représentés par les élus; réaliser les promesses faites aux électeur n'est pas l'objectif que poursuivent les élus*) qui organisent les cliques. A un seuil inférieur apparaît dans cette zone l'item B\* (*le fossé entre la société et la classe politique ne tend pas à se réduire*) et un peu plus à gauche, l'item F\* (*les partis politiques n'agissent pas conformément aux vœux des électeurs*). Enfin, complètement à gauche, le graphe est organisé à partir de l'item J\* (*Bien des hommes politiques ne sont pas animés par le souci du collectif*). A un seuil inférieur on trouve l'item K (*c'est le goût du pouvoir qui pousse à se présenter à des élections*).

De manière complètement séparée de cet ensemble de cliques on trouve des cliques plus centrées sur les institutions et les partis. Elles ont en commun les items E\* et M\* (*si les institutions se sont maintenues jusqu'ici, c'est que ce ne sont pas de bonnes institutions; réaliser les promesses faites aux électeurs n'est pas l'objectif que poursuivent*

*les élus*) auquel est associé soit J\* (souci du collectif) soit F\* (partis et voeux des électeurs). Enfin on trouve tout un ensemble de cliques isolées les unes des autres dont trois d'entre elles s'organisent autour des items I, J\* et A\*, c'est-à-dire I (*les partis sont devenus des organisations où la communication est bloquée*), A\* (*la constitution de la V république ne fonctionne pas bien*) et J\* (*bien des hommes politiques ne sont pas animés par le souci du service collectif*).



La vision en termes de cliques maximales de ce graphe montre bien qu'il n'y a pas, derrière ce questionnaire, d'axe unique. L'appréciation du champ politique existe sous une forme diversifiée: d'un côté à travers un Filtrant en zig-zag exprimant une appréciation ambiguë des personnalités politiques ou des comportements politiques des électeurs; de l'autre une appréciation des institutions. Celles-ci ne sont pas autonomisées car les hommes politiques sont toujours omniprésents.

#### 4 PLUS LOIN

L'intérêt de construire des graphes orientés est le dernier point abordé par l'article de Nicolas Roussiau et al. Là encore, comme dans le débat sur l'intérêt de calculer sur l'ensemble des corrélations ou sur celles concernées par un localité particulière du graphe (clique, cycle, co-cycle) il faut revenir aux principes théoriques que reflètent ces méthodes. Y-a-t-il un rapport entre structure causale et structure de similitude ? Dès son origine l'analyse de similitude repose sur un principe: la similitude est une opération symétrique (alors B ressemble à A) mais non transitive (si A ressemble à B et si B ressemble à C, il n'est pas sûr que A ressemble à C). Nous pensons même que la similitude est une des principales opérations cognitives portant sur les représentations sociales. Il est séduisant d'introduire dans la représentation sociale, des structures causales. Cela veut dire que certaines similitudes ne seraient pas symétriques et que certains items pourraient en entraîner d'autres. La corrélation ne serait à ce moment-là que le représentant d'une causalité et non pas d'une similitude. Mais pour avancer cela il faudrait avoir des données exprimant une causalité. Or ce n'est pas le cas de celles qui sont présentées ici. Elles permettent seulement de décrire comme nous l'avons montré au cours de la troisième conférence internationale des représentations sociales (Aix 1997, à paraître), des modèles collectifs que les personnes ou sujets peuvent actualiser. Le choix entre deux ou trois modèles crée les différences dans les réponses. L'analyse met en évidence que ces modèles organisent et structurent les réponses. Avant d'avancer des méthodes de traitement de données mettant en évidence des causalités il faut construire un modèle théorique des représentations sociales où la causalité prend place comme

mécanisme cognitif. La similitude est un modèle plus pauvre mais aussi plus adapté à l'état actuel de la théorie des représentations sociales.

### **BIBLIOGRAPHIE**

- Allport, F.H., 1934, The J-curve hypothesis of conforming behavior, *Journal of Social Psychology*, 5, pp 141-183.
- Flament, C., 1981, L'analyse de similitude: une technique pour les recherches sur les RS., *Cahiers de Psychologie Cognitive*, 1, pp 375-396.
- Flament, C., 1986, L'analyse de similitude: une technique pour les recherches sur les R.S., In W. Doise et A. Palmonari (éds), *L'étude des représentations sociales*, Neuchatel: Delachaux et Niestlé, pp 139-156.
- Vergés, P., 1995, Représentations sociales partagées, périphériques, indifférentes, d'une minorité: méthodes d'approche, *Les cahiers internationaux de psychologie sociale*, n°28, pp 77-95.

Pierre Vergés  
CRES  
3-5 Av. Pasteur  
13617 Aix en Provence  
France