

Ludovic Lebart

Marie Piron

Jean-François Steiner

# **La Sémiométrie**

**Essai de statistique structurale**



# **La Sémiométrie**

**Essai de statistique structurale**



## Avant-propos

Cet ouvrage s'adresse aussi bien à des statisticiens qu'à des psychosociologues, des spécialistes du marketing, de la communication, des linguistes, et plus généralement au grand public cultivé. La sémiométrie est en effet à la fois une curiosité philosophique, un outil du marketing et un champ d'expérimentation exceptionnel pour les statisticiens.

Dans ces diverses disciplines, l'ouvrage devrait intéresser un public très large allant des praticiens des bureaux d'études ou des instituts de sondage, aux étudiants, professeurs et chercheurs des universités ou des grandes écoles.

Plusieurs lectures sont possibles selon les connaissances du lecteur en mathématique et statistique, car des annexes et un glossaire, en fin d'ouvrage, permettent d'isoler les parties très techniques du texte, tout en en donnant une présentation assez simple, renvoyant sur des références accessibles et plus complètes.

L'ensemble doit beaucoup à des collaborations et des cadres de travail divers : au sein du département Economie Gestion et Sciences Humaines de l'Ecole Nationale Supérieure des Télécommunications, dirigé par Laurent Gille, et de l'URA 820 du Centre National de la Recherche Scientifique (Laboratoire de Traitement et Communication de l'Information, dirigé par Henri Maître) ; au sein de l'Institut de Recherche pour le Développement (IRD).

Nous sommes redevables de très nombreuses corrections, remarques, et suggestions à cinq personnes dont le travail de relecture et de critique fut en fait un véritable travail de collaboration. Il s'agit de Zysla Belliat et Hélène Haering, toutes deux membres du Comité Scientifique du Centre d'Etude des Supports de Publicité ; de Belaïd Ghermani de l'Université de Paris-Créteil ; de Françoise Potier de l'Institut National de Recherche sur les Transports et leur Sécurité ; de Jacqueline Sultan de la Société EMSIS.

Nous remercions également pour leurs remarques et observations pertinentes Christine Bellavoine, Etienne Brunet, Michel Grojnowski, Marie-Claude Lebart, Gérard Miltelberg, Armand Morgenzstern, Christine Tobin.

Nous restons évidemment seuls responsables des erreurs ou omissions qui pourraient subsister dans le texte.

Nous voudrions aussi dire toute notre gratitude à Jeannine Bouvet qui s'est livrée, avec une parfaite intelligence et le plus total désintéressement aux multiples travaux de saisie qui ont été à la base des différentes expériences décrites aux chapitres 3 et 4.

Nous sommes heureux d'adresser ici nos remerciements à Bruno Colin, Directeur des Opérations de la Société Taylor Nelson Sofres, notamment pour l'accès aux corpus de données sémiométriques.

Nos remerciements vont également à Jean Henriet, des éditions Dunod, ainsi qu'à Virginie Catoni, pour l'accueil qu'ils ont réservé à cet ouvrage.

L.L., M. P, J.-F. S.

Paris, 2003

# Sommaire

<b>Préambule</b>	1
<b>Introduction : Qu'est-ce que la sémiométrie ?</b>	5
<b>Chapitre 1</b>	
<b>Des mots, des individus... une structure</b>	9
<b>1.1 Des mots, un questionnaire</b>	10
<b>1.2 L'enquête représentative</b>	14
<b>1.3 L'analyse en composantes principales</b>	17
<b>1.4 Les axes principaux stables : « axes sémiométriques »</b>	19
1.4.1 Axe 1 : Un axe méthodologique de participation	19
1.4.2 Axe 2 : « Devoir / Plaisir »	21
1.4.3 Axe 3 : « Attachement / Détachement »	22
1.4.4 Le plan sémiométrique des axes 2 et 3	24
1.4.5 Axe 4 : « Esprit / Matière »	24
1.4.6 Axe 5 : « Cœur / Raison »	25
1.4.7 Axe 6 : « Humilité / Souveraineté »	26
1.4.8 Les axes suivants	27
<b>1.5 Les plans sémiométriques</b>	29
1.5.1 Les multiples facettes du « Plaisir »	29
1.5.2 Les multiples facettes du « Devoir »	35
1.5.3 Les dimensions sémantiques des mots eux-mêmes	36
<b>1.6 Conclusion</b>	37
<b>Chapitre 2</b>	
<b>Stabilité de la structure sémiométrique</b>	39
<b>2.1 Stabilité empirique</b>	40
2.1.1 Stabilité dans le temps	40
2.1.2 Stabilité par pays	41
2.1.3 Stabilité selon le sexe	51
2.1.4 Stabilité selon l'âge	53
<b>2.2 Intervalles de confiance des valeurs propres</b>	55
<b>2.3 Stabilité par rééchantillonnage</b>	58
2.3.1 Bootstrap partiel	59
2.3.2 Bootstrap total	59
<b>2.4 Stabilité structurelle</b>	64
<b>2.5 Stabilité vis-à-vis du codage</b>	67
<b>2.6 Analyse en facteurs communs et spécifiques</b>	69
<b>2.7 Conclusion</b>	71

<b>Chapitre 3</b>	
<b>Structures sémantiques et sémiométrie</b>	73
<b>3.1 Les « champs sémantiques internes »</b>	75
3.1.1 Représentation hiérarchique des proximités sémantiques	76
3.1.2 Représentation plane des proximités sémantiques	77
<b>3.2 Voisinage sémantique du questionnaire sémiométrique</b>	81
<b>3.3 Réseau sémantique des verbes français</b>	86
<b>Chapitre 4</b>	
<b>Choix spontané des mots</b>	93
<b>4.1 Mise en place de l'expérience</b>	94
<b>4.2 Premières explorations des réponses</b>	99
<b>4.3 Etude des axes principaux des réponses spontanées</b>	102
<b>4.4 Choix spontané et caractéristiques des répondants</b>	105
<b>4.5 Rapprochement entre sémiométrie et questions ouvertes</b>	110
<b>4.6 Conclusion</b>	116
<b>Chapitre 5</b>	
<b>Notations, participation, attitudes</b>	119
<b>5.1 Notation, participation et facteur de taille</b>	120
5.1.1 Les effets <i>notation</i>	120
5.1.2 Les effets <i>participation</i>	120
5.1.3 Les facteurs de taille	121
<b>5.2 Création et positionnement de variables de contrôle</b>	121
<b>5.3 Les notations dans quatre pays européens</b>	124
<b>5.4 Typologie à partir des seules variables de contrôle</b>	127
<b>5.5 Quels rapports entre les notes données aux mots, la façon de noter, et les caractéristiques des individus ?</b>	130
<b>5.6 Les plans de notation en Espagne, Grande Bretagne et Allemagne</b>	135
5.6.1 Effets <i>notation</i> en Espagne	135
5.6.2 Effets <i>notation</i> en Grande Bretagne	136
5.6.3 Effets <i>notation</i> en Allemagne	137
<b>5.7 Correction des notes par double centrage</b>	138
<b>5.8 Evolution des notations, en France, de 1990 à 2002</b>	142
<b>5.9 Conclusion</b>	145
<b>Chapitre 6</b>	
<b>Essais d'interprétation</b>	147
<b>6.1 Une première approche</b>	148
<b>6.2 Une interprétation psychanalytique</b>	152



<b>Chapitre 7</b>	
<b>Quelques applications de la sémiométrie</b>	159
7.1 Les mots caractéristiques des hommes et des femmes	161
7.2 Les mots caractéristiques de deux classes d'âge	162
7.3 Les grands amateurs de télévision	163
7.4 Les mots caractéristiques de deux classes de jeunes	165
7.5 Les mots caractéristiques de deux classes d'âge au delà de 60 ans	166
7.6 Hommes et femmes ayant attribué la note 7 au mot Dieu	168
7.7 Evolution des valeurs des Français entre 1996 et 1998	170
7.8 Conclusion	173
<b>Conclusion</b>	175
<b>Annexe 1</b>	
<b>Quelques éléments de statistique multidimensionnelle</b>	183
<b>A1.1 Rappel des principes des méthodes exploratoires         multidimensionnelles</b>	184
A1.1.1 Représentation géométrique et nuages de points	184
A1.1.2 Principe et méthodes d'analyse	185
<b>A1.2 Les méthodes factorielles : aspects techniques</b>	187
<b>A1.3 L'Analyse en Composantes Principales :         aspects techniques</b>	190
A1.3.1 Interprétations géométriques	190
A1.3.2 Problème d'échelle de mesure et transformation des données	190
A1.3.3 Analyse du nuage des répondants	191
A1.3.4 Analyse du nuage des variables (mots)	192
<b>A1.4 L'Analyse des correspondances</b>	194
<b>A1.5 L'analyse logarithmique</b>	195
<b>A1.6 L'analyse factorielle en facteurs communs et spécifiques</b>	196
<b>A1.7 Méthodes de classification hiérarchique</b>	197
<b>A1.8 Les cartes auto-organisées de Kohonen</b>	199
<b>A1.9 Outils de validation</b>	200
A1.9.1 Qu'est-ce qu'une valeur-test ?	201
A1.9.2 Le problème des comparaisons multiples	203
A1.9.3 Utilité des éléments supplémentaires	204
A1.9.4 Intervalles de confiance d'Anderson	205
A1.9.5. Les techniques de <i>bootstrap</i>	205
<b>Annexe 2</b>	
<b>Quelques éléments et résultats d'analyse</b>	211
<b>A2.1 Dictionnaire multilingue des mots</b>	211

<b>A2.2 Moyenne des notes</b>	216
<b>A2.3 Corrélations des mots par axe</b>	218
<b>Glossaire</b>	221
<b>Bibliographie</b>	225

## Préambule

### *A l'origine de la Sémiométrie...*

La Sémiométrie est née d'une expérience que tout homme a faite au moins une fois dans sa vie, d'une manière consciente et lucide ou si fugace que l'esprit n'en a gardé aucune trace mémorable.

Je veux parler de cet instant où, soudain, pour quelque raison mystérieuse, nous cessons de communiquer avec les êtres et les objets qui nous entourent, où ils cessent de nous parler, où ils ne nous disent plus rien, où nous nous retrouvons tout seuls au milieu d'un monde minéral peuplé d'objets morts, et dans lequel, les sens glacés, l'esprit figé, nous ne savons plus nous orienter. Quelques instants ou quelques jours auparavant – le phénomène se déclenche d'une manière plus ou moins brutale et sa durée est extrêmement variable – le monde était plein de vie, de sons, de couleurs, de saveurs, d'odeurs, de sensations agréables ou désagréables, d'objets et d'êtres qui nous attiraient, d'autres qui nous repoussaient, que nous aimions ou que nous haïssions, mais que nous pouvions nommer et dont le nom signifiait quelque chose pour nous, était doué d'un pouvoir évocateur qui faisait naître, à sa seule énonciation, un cortège de sensations qui en étaient comme la signature : un ami était quelque chose de chaud et de précieux, de rare aussi, de solide, très solide, et pourtant, en même temps, de très fragile ; une maison était aussi quelque chose de chaud, mais d'une autre chaleur ; alors que la chaleur de l'ami irradiait à partir d'un point situé au milieu de la poitrine, celle de la maison semblait plutôt descendre de quelque

lieu inconnu et extérieur à nous et nous pénétrer lentement de bonheur ; la chaleur de l'ami était aussi plus intense, plus violente que celle de la maison ; c'était une chaleur rouge alors que celle de la maison tirait plutôt sur le bleu. Le mot *guerre* aussi signifiait quelque chose ; c'était un tout autre monde de sensations qui semblaient appartenir à une zone qui se situait aux antipodes de celle à laquelle appartenait le mot *maison*, mais qui semblait avoir un territoire commun avec le mot *ami*. Le mot *arbre* aussi avait une signification et le mot *rivière* et le mot *chemin* et le mot *école* et le mot *été* et le mot *forêt* et le mot *tristesse* et même le mot *ennui*, et jusqu'au mot *désespoir*.

Bref, grâce aux signes qui permettent de le décrire, le monde avait un sens ; il nous était intelligible ; nous savions nous y orienter ; il était notre milieu naturel, si naturel même que la pensée ne nous effleurait pas que nous pourrions en être exclus autrement que par la mort. Nous étions si accoutumés à le retrouver chaque matin à notre réveil, toujours identique à lui-même – à l'exception de quelques objets qui de temps en temps, migraient d'une zone dans une autre –, que nous n'avions pas réfléchi à cette évidence – si forte, pourtant, dès lors qu'on la conçoit – que le sens n'est pas consubstantiel au monde qui nous entoure, qu'il n'en est qu'une propriété que nous lui attribuons, qu'une autre réalité que nous projetons sur lui. Il avait fallu cette expérience de perte du sens – dans la triple acception du mot : sensation, signification et direction – pour nous faire appréhender ce fait bien banal que le sens est un phénomène autonome et indépendant des objets à travers lesquels il s'incarne, qu'il est une réalité, au moins psychique sinon physique, et donc, qu'il doit pouvoir être étudié indépendamment des objets auquel il sert d'attribut et par la médiation desquels on le perçoit généralement.

Mais cette première expérience en contenait une autre, qui en était comme la suite logique : non seulement il y avait DU sens, mais il y avait aussi DES sens. Ainsi le mot *château*, non seulement pouvait être doué (ou dénué) de sens mais, dès lors qu'il possédait cette propriété, il en découlait nécessairement qu'il possédait aussi UN (ou plusieurs) sens.

Et cette deuxième expérience, à son tour, en contenait implicitement une troisième : ces différents sens révélés par les mots étaient plus ou moins proches les uns des autres. Ainsi le sens du mot *château* était intuitivement plus proche de celui du mot *maison* que de celui du mot *mathématique*, mais plus proche encore de celui du mot *palais*. Le sens du mot *maison* semblait équidistant de celui des mots *château* et *palais*. Le sens du mot *mathématique* était si éloigné de celui de ces trois

derniers qu'il était difficile de dire s'il était plus proche de l'un d'entre eux, de la même manière que trois points très proches semblent pareillement distants d'un quatrième, si celui-ci est situé à une grande distance ; quoique, à y regarder de plus près, le sens du mot *château* paraissait très légèrement moins éloigné de celui du mot *mathématique* que ne l'était celui du mot *maison*. Comme pour le sens des autres mots, ce n'était qu'une intuition, mais si ténue, si subtile, si insaisissable qu'elle s'évanouissait dès que nous essayions de la raisonner, et qui pourtant renaissait presque instantanément, comme quelque tache indélébile, dès que nous détournions d'elle la lumière trop crue de notre attention. Ces intuitions qui nous faisaient attribuer une distance de sens entre les mots possédaient donc un fondement ; elles étaient l'expression d'une réalité qui, pour apparemment inaccessible à un instrument d'observation objectif, n'en était pas moins stable et, donc, théoriquement mesurable. On pouvait donc imaginer un espace qui contiendrait l'ensemble des mots d'une langue, disposés les uns par rapport aux autres en fonction de la distance séparant leur sens. Ce serait une sorte de dictionnaire géographique dans lequel les mots ne seraient pas rangés dans l'ordre alphabétique – ce qui, pour pratique qu'il soit, n'en est pas moins un classement relativement arbitraire – mais disposés, les uns par rapport aux autres, en fonction de leur sens, un dictionnaire organique dans lequel chaque mot serait défini (mathématiquement) par sa position dans l'ensemble de tous les autres. Par référence à Ferdinand de Saussure qui, imaginant une science des signes, avait proposé de l'appeler sémiologie, je baptisai mon entreprise, sémiométrie<sup>1</sup>.

\*

\* \*

*J'éprouve une très grande gratitude à l'égard des personnes qui m'ont aidé dans cette entreprise. Je voudrais d'abord citer le nom de Philippe Court dont l'amitié indéfectible et une foi inébranlable m'ont soutenu dès le premier instant. Je voudrais aussi citer celui d'Eric Stemmelen, à l'époque ingénieur au Commissariat à l'Energie Atomique, qui m'a appris les premiers rudiments de l'analyse de données, avant de diriger le département que la Sofrès avait créé pour développer commercialement les*

---

1. L'idée qu'il existe un espace du sens pouvant être modélisé revient à Charles E. Osgood. C'est lui, en effet, qui, dans les années 50, imagina le premier un dispositif destiné à reconstruire cet espace. Ce dispositif diffère radicalement de celui mis en oeuvre par la Sémiométrie et j'ignorais tout des travaux d'Osgood lorsque je conçus celui-ci. Mais il convenait de citer le nom de celui qui avait introduit ce concept si riche d'espace de sens, cf. Osgood (1965), Osgood *et al* (1957).

*applications de la sémiométrie. Dans l'ordre chronologique, il me faut ensuite citer ceux de Georges Levitte et de Léon Szary qui a écrit le premier programme grâce auquel j'ai pu procéder aux premières expérimentations.*

*S'il ne fut pas mon premier mentor, Emeric Deutsch, à l'époque l'un des dirigeants de la Sofrès, fut certainement le plus décisif, grâce à la profondeur de ses connaissances tant en psychanalyse et en linguistique que dans le domaine des études de marché.*

*Je dois aussi remercier la Sofrès, son Président, Pierre Weill qui a mis à ma disposition des moyens dont peu de chercheurs disposent, son Directeur-Général, Henri Wallard, dont l'amitié me fut bien précieuse, Jean-Michel Portier, Jean-Paul Aimetti, Joël Boillot, Mireille Gettler-Summa, Jean-Pierre Dreyfus, Valérie Morrison, et tant d'autres.*

Jean-François Steiner

## **Introduction :** **Qu'est-ce que la sémiométrie ?**

A la question « Qu'est-ce que la sémiométrie ? », il est facile de répondre naïvement : « C'est une longue liste de mots auxquels on fait attribuer une note (selon leur caractère plus ou moins agréable ou désagréable) par des milliers de gens ».

Tentons quand même des caractérisations plus académiques :

- la sémiométrie est une technique de description de certains types de liens sémantiques entre mots ;
- la sémiométrie est un instrument de recherche plus fondamentale permettant d'approcher (avec prudence et circonspection) des notions telles que l'« inconscient collectif » ou individuel, permettant aussi de mettre en évidence et de valider des traits structuraux (dont la finesse laisse parfois pantois...) ;
- la sémiométrie est un outil destiné à décrire des systèmes de valeurs et des styles de vie, en vue d'études psycho-sociologiques ou d'applications en marketing.

La sémiométrie est en fait un peu tout cela, et, c'est ce que cet ouvrage propose de montrer.

La première définition « naïve » est l'évocation d'un protocole expérimental élaboré et rigoureux, auquel sera d'ailleurs dévolu le premier chapitre de cet ouvrage. Cette phase expérimentale qui comporte de nombreuses enquêtes répétées dans le temps et dans

l'espace (notamment dans plusieurs pays occidentaux) constitue la base de toute la méthodologie sémiométrique, base qui sera validée par des traitements statistiques puissants, mais aveugles, c'est-à-dire indépendants de tout cadre interprétatif. Notons que, au niveau de la présentation, dans ce premier chapitre comme dans ceux qui vont suivre, les développements techniques seront reportés en annexe de façon à permettre une lecture plus directe du contenu, des résultats, et de la méthodologie générale, sans heurt ni digression.

Le thème de la stabilité de structures (stabilité dans le temps et l'espace) qui fait l'objet du deuxième chapitre, est central dans la problématique de la sémiométrie. Cette stabilité est établie à partir de la répétition d'enquêtes indépendantes, mais aussi à partir d'épreuves de validité plus techniques (méthodes statistiques dites de rééchantillonnage). C'est cette stabilité qui permet d'affirmer la réalité de certains traits structuraux observés et qui donne aux cartes sémiométriques, largement utilisées lors des applications, leur valeur de grilles de repérage.

La stabilité des structures pour un questionnaire donné (comportant le choix d'une liste de mots) étant établie dans différents pays et à différents moments, il reste le problème du caractère intrinsèque de cette structure, c'est-à-dire de son indépendance vis-à-vis du questionnaire lui-même, et non pas seulement vis-à-vis des populations auxquelles ce questionnaire est soumis.

Les trois chapitres suivants évoquent nos recherches méthodologiques autour de ce thème. Pour mieux comprendre la nature des faits statistiques observés, diverses expériences ont été réalisées : étude de certains réseaux sémantiques attachés aux mots choisis, puis étude de réseaux sémantiques plus généraux (chapitre 3), questionnaire « ouvert » laissant la personne interrogée choisir spontanément, sans liste préalable, les mots qu'elle trouve agréables ou désagréables (chapitre 4). Ces expériences sont d'ailleurs l'occasion d'aborder d'autres directions de recherche suscitées par l'outil sémiométrique et la structure multidimensionnelle qu'il révèle. Le cinquième chapitre est dévolu au problème technique, souvent éludé, du lien entre l'attitude de la personne interrogée par rapport au questionnaire ou à l'enquête et le contenu même du questionnaire ou la finalité de l'enquête.

Légèrement en retrait par rapport aux considérations empiriques et techniques précédentes, un essai d'interprétation de la partie stable des



structures sémiométriques est proposé au sixième chapitre. Quel que soit le degré d'adhésion du lecteur aux thèses présentées, il ne sera probablement pas insensible aux étonnantes convergences mentionnées.

Enfin, au cours du septième chapitre, dévolu aux applications de l'outil sémiométrique, nous serons amenés à présenter des applications pratiques ou pragmatiques (en général commerciales : média-planning, positionnement ou caractérisation de marques), des applications à des problèmes plus généraux (évolution des valeurs, caractérisation de groupes socio-démographiques particuliers).

Le chapitre 8 est en fait une conclusion, qui ne peut être que provisoire dans le cadre d'un travail aussi périlleusement pluridisciplinaire ; c'est aussi une invitation à des recherches futures, dans un domaine que nous proposons d'appeler « statistique structurale ».

Viennent enfin les annexes, austères entrepôts de matériaux, de résultats, de rappels techniques, où le lecteur motivé par les chapitres précédents pourra trouver des réponses à beaucoup de questions soulevées au cours de son voyage dans cet espace multidimensionnel.

