

## CHAPITRE 2

### Stabilité de la structure sémiométrique

Nous allons montrer dans ce chapitre que la structure mise en évidence au chapitre premier est stable, c'est-à-dire n'est ni un accident statistique, ni un artefact des méthodes utilisées. Cette structure est un ensemble de six axes principaux bien identifiables, auxquels nous avons donné des noms conventionnels, mais que tout un chacun peut reconnaître, indépendamment de cette dénomination, par les mots qui leurs sont associés.

Dans un premier temps, la stabilité sera éprouvée empiriquement sur différents échantillons recueillis à des moments différents pour un même pays (la France); il s'agit donc de *stabilité dans le temps*. Puis des échantillons relatifs à des pays différents seront analysés; on parlera alors de *stabilité dans l'espace*. Enfin, les axes sémiométriques seront recalculés dans des sous-populations particulières à l'intérieur d'un même pays, étudiant séparément les hommes et les femmes, puis différentes classes d'âge. On parlera alors de *stabilité interne*. Quelques exceptions et singularités seront néanmoins observées. Elles seront limitées, mais intéressantes à interpréter, comme peuvent l'être de petites déviations par rapport à un modèle ou à une norme. On constatera ainsi quelques différences subtiles entre Europe du nord et du sud. Les axes sont stables, mais leur ordre d'apparition peut varier. On constatera aussi, au chapitre 5 cette fois, la lente dérive de la population française au sein de cette structure au cours des douze dernières années.

Dans un second temps, cette stabilité sera testée de façon plus technique, utilisant les outils de la validation statistique et de la simulation. Les

échantillons seront perturbés par plusieurs variantes d'une méthode de *rééchantillonnage* connue sous le nom de *bootstrap*<sup>1</sup> pour mettre à l'épreuve la structure sémiométrique. Celle-ci sera, disons-le d'emblée, largement insensible à ces perturbations.

Evidemment, ce chapitre va se livrer à un exercice peu captivant pour un lecteur en quête d'aventures inédites... puisqu'il s'agit en somme de montrer que l'on trouve toujours la même chose, dans le temps, dans l'espace, voire à l'intérieur même de certaines sous-catégories d'individus. Le lecteur qui nous ferait confiance sur la surprenante stabilité de la structure sémiométrique peut donc passer au chapitre suivant, et simplement revenir consulter celui-ci si un doute le saisit lors de ses cheminements à venir dans la richesse et la complexité de l'espace des mots.

## 2.1 Stabilité empirique par sous-population

### 2.1.1 Stabilité dans le temps

Le questionnaire sémiométrique a été passé en France durant six années réparties sur une période de dix ans (1990, 1996, 1998, 1999, 2001 et 2002) en vue d'applications marketing<sup>2</sup> et pour suivre certains aspects de l'évolution des valeurs des Français. La comparaison de ces vagues successives d'enquêtes obtenues à partir d'échantillons indépendants met en évidence une grande stabilité des structures (*cf.* tableau 2.1).

Mais, si les structures sémiométriques ne changent pas en elles-mêmes, on observe cependant de légères différences au niveau du positionnement de certaines catégories. Des modifications de faible amplitude peuvent être relevées selon le sexe, l'âge mais surtout l'activité. Ainsi, les exploitants-agricoles expriment de plus en plus depuis ces dix dernières années des valeurs proches du pôle « Détachement » (un des pôles de l'axe 3 : « Attachement / Détachement ») et semblent moins caractérisés par le pôle « Devoir » de l'axe 2 : « Devoir / Plaisir ». Les artisans-commerçants et les ouvriers se rapprochent en 1998 du pôle « Plaisir » de ce même axe, caractérisé avant cette date surtout par les cadres et les professions intermédiaires. Ces déplacements de catégories dans l'espace sémiométrique, faibles mais significatifs statistiquement, demandent pour être interprétés un travail pluridisciplinaire impliquant aussi bien des statisticiens que des sociologues et des économistes. Un tel travail sort du

---

1. Le principe de la méthode « *bootstrap* » est esquissé en annexe A1.9.5.

2. Des exemples d'application font l'objet du chapitre 7.

cadre de la présentation plus spécifiquement méthodologique de cet ouvrage.

Nous verrons au chapitre 5 que la manière de noter rend compte d'une autre évolution significative qu'il n'est pas possible de déceler ici sur la simple comparaison d'analyses effectuées indépendamment sur chaque échantillon. Nous renvoyons pour une analyse de l'évolution des valeurs des Français au paragraphe 7.7 du chapitre 7 consacré aux comparaisons diachroniques entre des échantillons d'une même population interrogés à des moments différents.

Bien que l'on observe des mouvements significatifs de catégories dans les plans sémiométriques, les positions des mots dans ces plans restent stables, comme l'attestent les tableaux 2.1 à 2.6 qui décrivent systématiquement les évolutions des axes 1 à 6 pour les quatre années 2002, 1999, 1996, 1990. Dans cette série de tableaux, chaque axe, pour chaque année, est décrit par ses 20 mots les plus extrêmes, c'est-à-dire par les dix mots qui lui sont le plus corrélés positivement et les dix mots qui lui sont le plus corrélés négativement.

### **2.1.2 Stabilité par pays**

Le questionnaire sémiométrique a été soumis à des échantillons appartenant à plusieurs pays européens<sup>1</sup>, toujours en vue d'applications de type marketing. Une des inconnues, dans cette opération d'envergure, était l'effet de la traduction du questionnaire sur les résultats. Le tableau A2.1 présenté en annexe A2 donne les correspondances entre mots dans les cinq langues concernées. Il montre à la fois l'étendue et les limites des difficultés rencontrées. Beaucoup de mots ne posent pas de gros problèmes et d'autres conduisent à faire des choix parfois difficiles<sup>2</sup>.

---

1. France, Grande Bretagne, Allemagne, Italie, Espagne, Finlande, Norvège, ....

2. Ainsi, *Dieu*, *L'eau*, *Un arbre*, *Métallique*, *Un nid*, ne posent pas de gros problèmes, en première approximation, et pour les pays concernés. En revanche, *Argent*, *Réfléchir*, *un Prêtre*, sont des mots dont la traduction n'est pas immédiate (*Silver* ou *Money* pour *Argent* en Anglais ?) et nécessitent un choix raisonné, voire une convention arbitraire. Mais comme on le verra plus loin, les résultats seront également stables par *perturbation* de la liste de mots, et la traduction peut être considérée comme une opération entraînant une perturbation dans des limites tolérables.

**Tableau 2.1 :**  
**Mots caractérisant les « axes 1 » de 1990 à 2002**  
**« Axe de participation »**

axe 1 : 2002	axe 1 : 1999	axe 1 : 1996	axe1 : 1990
courage	courage	courage	protéger
dynamique	efficace	protéger	pureté
douceur	dynamique	efficace	courage
efficace	précision	dynamique	honnête
gaieté	récompense	robuste	récompense
respect	élégance	honnête	efficace
récompense	robuste	politesse	admirer
honnête	charitable	pureté	dynamique
confiance	confiance	élégance	respect
élégance	protéger	récompense	politesse
zone centrale			
faute	noir	orage	doute
rompre	mort	doute	mort
attaquer	danger	rompre	vide
mort	rompre	guerre	faute
danger	faute	vide	danger
vide	vide	faute	guerre
angoisse	angoisse	danger	angoisse
guerre	désordre	angoisse	rompre
trahir	guerre	désordre	désordre
désordre	trahir	trahir	trahir

**Tableau 2.2 :**  
**Mots caractérisant les « axes 2 » de 1990 à 2002**  
**« Devoir / Plaisir »**

axe 2 : 2002	axe 2 : 1999	axe 2 : 1996	axe 2 : 1990
sensuel	sensuel	sensuel	sensuel
rêver	rêver	rêver	aventurier
voluptueux	légèreté	aventurier	rêver
île	sauvage	nudité	sauvage
légèreté	aventurier	original	original
moelleux	île	voluptueux	nudité
original	original	désir	bohème
océan	voluptueux	bohème	légèreté
sauvage	nudité	île	voluptueux
nudité	charnel	émotion	orage
zone centrale			
règle	économiser	interdire	prêtre
rigide	règle	punir	loi
prêtre	loi	règle	honneur
élite	industrie	industrie	interdire
soldat	prêtre	économiser	économiser
industrie	soldat	morale	soldat
patrie	obeir	soldat	patrie
obéir	patrie	patrie	discipline
morale	morale	discipline	morale
discipline	discipline	obéir	obéir

Exemple de lecture du tableau 2.1 : Pour l'axe 1, en 2002, le mot le plus corrélé positivement à l'axe est le mot courage, suivi par dynamique, douceur, efficace, etc. Le mot le plus corrélé négativement à ce même axe est le mot désordre, suivi de trahir, guerre, angoisse, etc.

**Tableau 2.3 :**  
**Mots caractérisant les « axes 3 » de 1990 à 2002**  
**« Attachement / Détachement »**

axe 3 : 2002	axe 3 : 1999	axe 3 : 1996	axe 3 : 1990
danger	danger	danger	danger
orage	mort	orage	doute
révolte	punir	mort	angoisse
critiquer	vide	désert	rompre
mort	rompre	labyrinthe	cri
attaquer	attaquer	révolte	vide
cri	angoisse	rompre	labyrinthe
feu	labyrinthe	critiquer	mort
détachement	critiquer	cri	orage
vide	révolte	angoisse	méfiance
zone centrale			
fidélité	fleur	mariage	argent
richesse	or	gratuit	richesse
mariage	gaîté	hériter	récompense
récompense	maison	récompense	caresse
cadeau	récompense	or	gaîté
maison	famille	argent	maison
argent	mariage	richesse	or
confort	confort	cadeau	confort
famille	cadeau	bijou	cadeau
bijou	bijou	confort	bijou

**Tableau 2.4 :**  
**Mots caractérisant les « axes 4 » de 1990 à 2002**  
**« Esprit / Matière »**

axe 4 : 2002	axe 4 : 1999	axe 4 : 1996	axe 4 : 1990
richesse	richesse	puissance	richesse
puissance	puissance	argent	puissance
or	or	richesse	argent
gloire	argent	or	or
vitesse	gloire	gloire	gloire
argent	conquérir	ruse	conquérir
ambition	vitesse	vitesse	vitesse
conquérir	commander	ambition	propriété
ruse	hériter	conquérir	ambition
séduire	ambition	fusil	hériter
zone centrale			
paix	pardon	école	école
écrire	enseigner	enseigner	recueillement
enseigner	écrire	art	pardon
réfléchir	arbre	pardon	âme
école	étranger	charitable	fleur
arbre	école	poésie	charitable
théâtre	poésie	recueillement	art
poésie	théâtre	nid	arbre
art	art	arbre	poésie
livre	livre	livre	livre

**Tableau 2.5 :**  
**Mots caractérisant les « axes 5 » de 1990 à 2002**  
**« Cœur / Raison »**

axe 5 : 2002	axe 5 : 1999	axe 5 : 1996	axe 5 : 1990
dieu	dieu	dieu	dieu
foi	foi	foi	foi
prêtre	âme	âme	prêtre
sacré	recueillement	prêtre	âme
âme	sacré	sacré	éternel
recueillement	prêtre	éternel	sacre
créateur	créateur	cérémonie	recueillement
éternel	éternel	recueillement	bijou
infini	noble	souverain	cérémonie
souverain	souverain	poésie	noble
zone centrale			
confort	volontaire	construire	industrie
réfléchir	produire	utilitaire	volontaire
matériel	maîtriser	astucieux	matériel
maîtriser	utilitaire	produire	robuste
volontaire	concret	réfléchir	maîtriser
utilitaire	astucieux	maîtriser	efficace
concret	construire	efficace	précision
logique	dynamique	précision	produire
précision	précision	concret	concret
efficace	efficace	logique	logique

**Tableau 2.6 :**  
**Mots caractérisant les « axes 6 » de 1990 à 2002**  
**« Humilité / Souveraineté »**

axe 6 : 2002	axe 6 : 1999	axe 6 : 1996	axe 6 : 1990
créateur	bâtitseur	noble	noble
art	créateur	créateur	élite
audace	maîtriser	art	commander
robuste	art	élite	audace
inventeur	robuste	dieu	dieu
efficace	concret	audace	créateur
bâtitseur	inventeur	souverain	art
précision	audace	foi	absolu
voluptueux	noble	âme	bâtitseur
noble	efficace	robuste	maîtriser
zone centrale			
obéir	cri	enfance	enfance
féconder	interdire	rigide	famille
rigide	obéir	obéir	tendresse
faute	mariage	doute	angoisse
vide	famille	mariage	obéir
mariage	tendresse	interdire	rigide
méfiance	méfiance	famille	faute
angoisse	doute	faute	doute
famille	faute	méfiance	méfiance
naissance	naissance	naissance	naissance

Les résultats présentés concernent six de ces pays : la France, la Grande Bretagne, l'Allemagne, l'Italie, l'Espagne, le Canada (Québec). La stabilité des structures peut être vérifiée à la lecture des figures 2.1 à 2.5.

On relève cependant certaines singularités qui, sans remettre en cause la stabilité d'ensemble, la nuancent, mettant en évidence des particularités culturelles sur lesquelles il est intéressant de s'arrêter.

La première de ces singularités concerne l'interversion des axes 2 et 3, « Devoir / Plaisir » et « Attachement / Détachement », entre les pays du sud de l'Europe et ceux du nord : pour les Espagnols, les Français et les Italiens, l'axe « Devoir / Plaisir » apparaît avant l'axe « Attachement / Détachement ». A l'inverse, pour les Allemands, les Anglais et les Canadiens, c'est la dimension « Attachement / Détachement » qui apparaît en seconde position, et qui précède donc l'axe « Devoir / Plaisir » occupant maintenant la troisième position<sup>1</sup>. Ce ne sont pas les axes qui sont en cause, mais leur hiérarchie<sup>2</sup>.

Le chapitre 5 montrera que cette interversion d'axes peut être imputable, du moins en partie, aux effets de notation qui caractérisent les premier et troisième axes et qui traduisent des différences culturelles d'attitude par rapport à l'enquête et/ou son questionnaire ainsi que des différences de comportement vis-à-vis d'une échelle de notes.

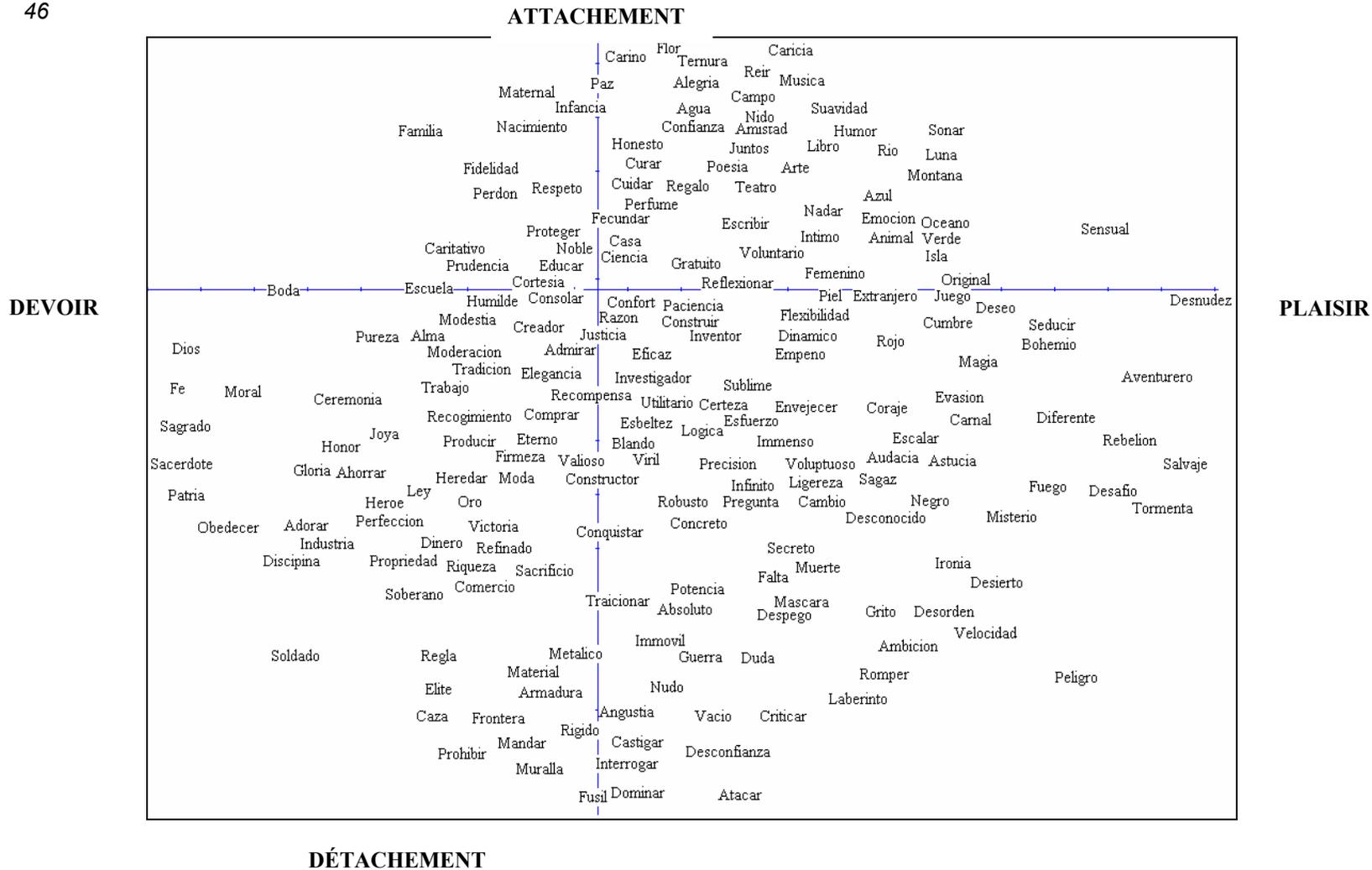
Cette singularité se retrouve également dans les analyses que l'on peut faire à l'intérieur de certaines classes d'âges. Ce point sera repris lorsque seront présentées ces analyses (section 2.1.4).

La deuxième singularité concerne l'axe 4 appelé conventionnellement « Esprit / Matière ». Si, du côté du pôle « Matière », on retrouve, pour les cinq pays considérés, un grand nombre de mots communs tels que *Or*, *Argent*, *Bijou* (sauf Canada), *Richesse*, *Conquérir*, *Victoire*, *Séduire*, etc., qui expriment très clairement la recherche d'une jouissance à travers la conquête et la possession de bien matériels, les mots qui s'y opposent à l'autre extrémité de l'axe diffèrent assez nettement selon les pays.

---

1. Ce résultat est confirmé par les axes de la Norvège, de la Finlande et de la République Tchèque, dont il n'est pas rendu compte ici.

2. On peut donc dire que, pour décrire l'ensemble des corrélations entre mots, l'opposition « Devoir / Plaisir » est plus importante que l'opposition « Attachement / Détachement » dans le sud de l'Europe que dans le nord de l'Europe.



**ATTACHEMENT**

**DEVOIR**

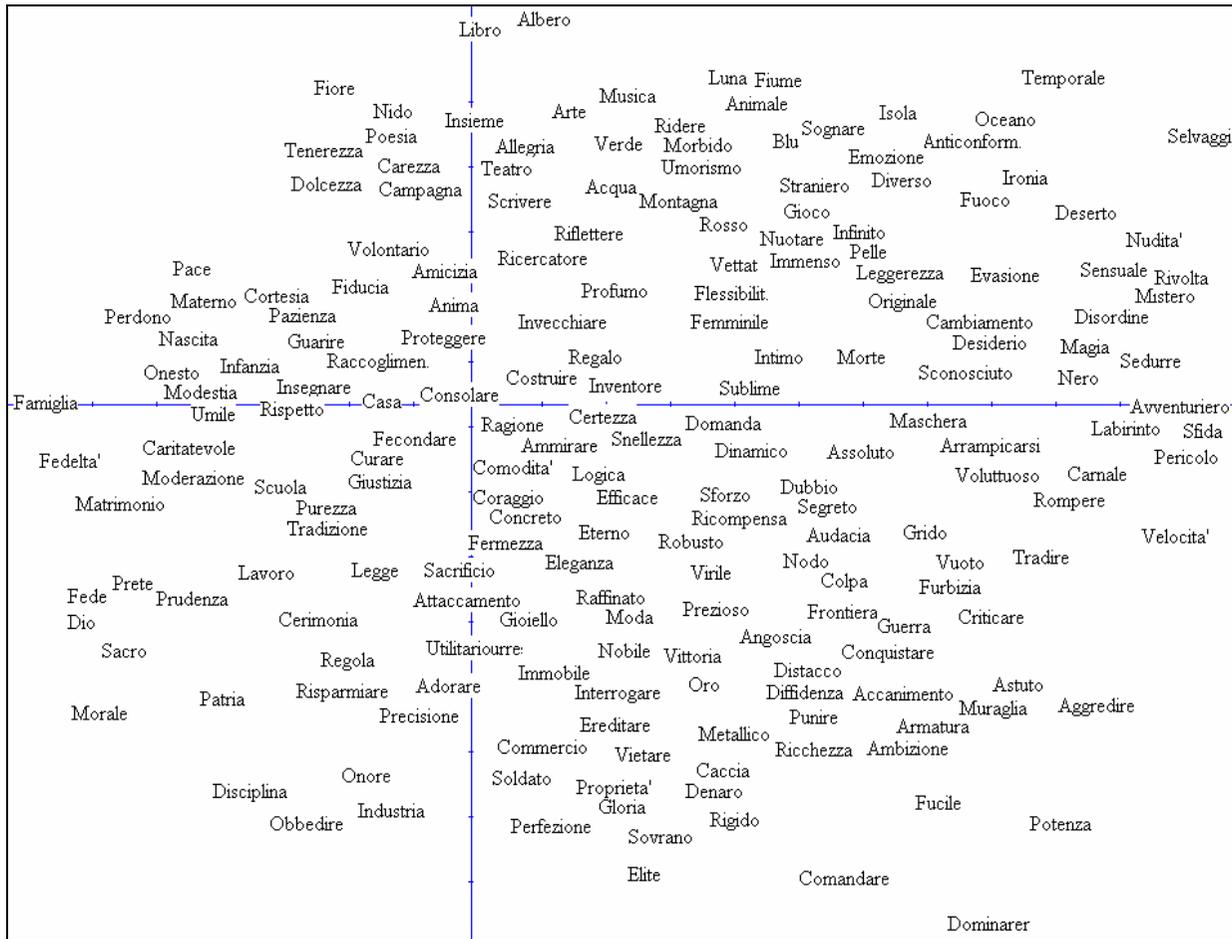
**PLAISIR**

**DÉTACHEMENT**

**Figure 2.1 : Plan sémiotique (2,3) - Espagne**

ATTACHEMENT

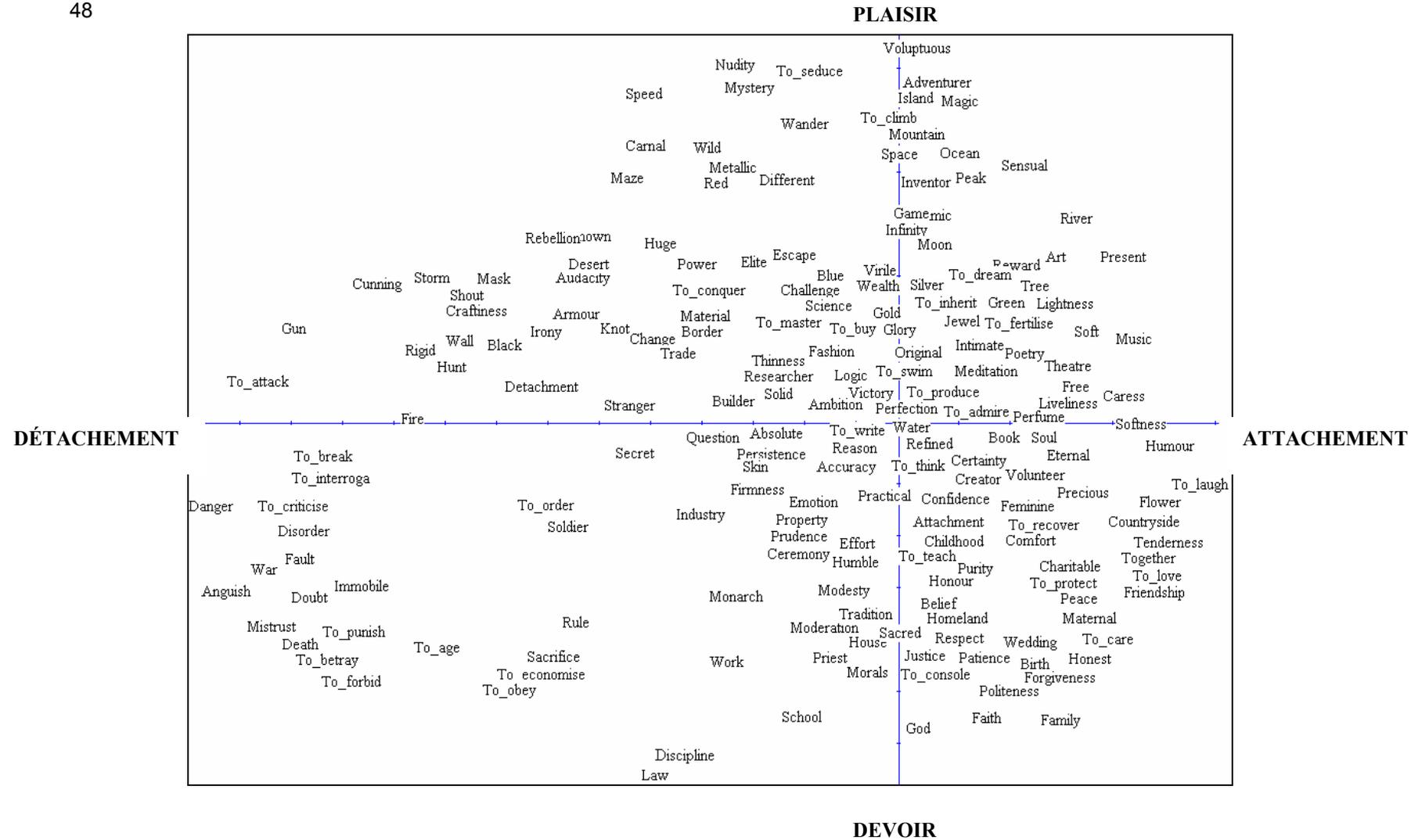
DEVOIR



PLAISIR

DÉTACHEMENT

Figure 2.2 : Plan sémiométrique (2,3) - Italie





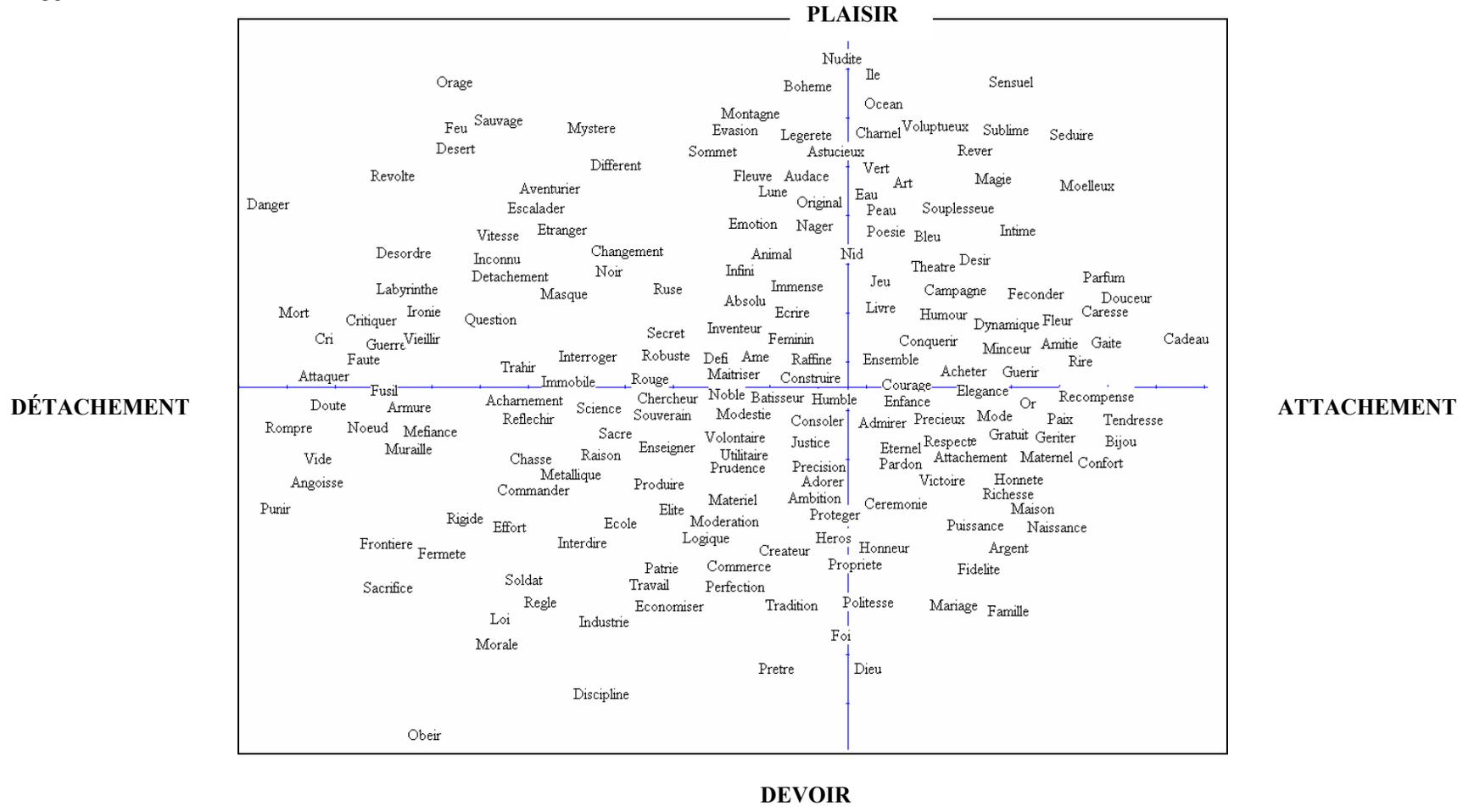


Figure 2.5 : Plan sémiométrique (2,3) - Canada

On peut diviser ceux-ci en trois groupes : la France où la notion d'« Esprit » ou encore de « Sublimation » est avant tout – mais non exclusivement – exprimée par l'ensemble des mots du questionnaire évoquant les activités artistiques (*Livre, Art, Théâtre et Poésie*), le Canada mais surtout la Grande-Bretagne où cette notion est principalement rendue par des mots liés à la nature (*Arbre, Montagne, Fleuve, Océan, Lune et Vert*), et l'Espagne et l'Italie où apparaissent des mots comme *Mort, Sacrifice, Humble, Faute* (sauf Espagne), *Doute*, qui sont complètement absents dans les autres pays ; l'Allemagne présentant une synthèse de ces trois groupes<sup>1</sup>.

### **2.1.3 Stabilité selon le sexe**

Nous avons vu au cours du chapitre précédent que les axes 3 (intitulé conventionnel : « Attachement / Détachement ») et 5 (intitulé conventionnel : « Cœur / Raison »), obtenus à partir d'analyses de la population globale, opposaient des populations féminines et masculines. Si la variable « sexe » est vraiment la variable explicative cachée responsable des grandes oppositions observées, les analyses réalisées uniquement auprès d'un échantillon de femmes ou uniquement auprès d'un échantillon d'hommes ne doivent plus faire apparaître d'axes de ce type.

Or, l'ensemble des structures sémiométriques restent, à quelques nuances près, conservées lors d'analyses ne portant que sur des hommes ou sur des femmes. Il y a là une stabilité et une permanence assez inattendues. La structure est la même pour les hommes et pour les femmes, mais certains des traits de cette structure peuvent séparer de façon significative les hommes des femmes<sup>2</sup>.

---

1. On pourrait, à propos de tels résultats, faire l'hypothèse que des particularités culturelles peuvent être à l'origine des différences observées, les Français opposant l'art à l'argent, les Anglais, l'or de la City à la campagne anglaise, et les habitants des pays du sud opposant alors l'or à la mort. Mais, conscients de nos limites dans un domaine aussi subtil, nous laisserons ce soin à nos lecteurs et aux spécialistes concernés.

2. Ce phénomène d'apparition d'une même structure à des échelles différentes évoque irrésistiblement les structures dites *fractales*. Ce phénomène, tout comme l'analogie qui y est attachée, fut déjà remarqué et mentionné par Jean-Pierre Pagès, dans le cadre des enquêtes Agoramétrie. L'analogie avec le monde fractal est d'autant plus prégnante qu'elle se prolonge, de subdivision en subdivision de l'échantillon, jusqu'à l'individu, qui porterait en lui une structure du même type. Nous reviendrons à diverses reprises sur ce thème, et ce jusqu'au moment où il s'agira pour nous de conclure et d'évoquer problèmes ouverts et recherches futures.

**Tableau 2.7 : Confrontation, axe par axe, des analyses réalisées séparément sur les hommes et sur les femmes**

axe 2 : « Devoir / Plaisir »		axe 3 : « Attachement / Détachement »		axe 4 : « Esprit / Matière »		axe 5 : « Cœur / Raison »		axe 6 : « Humilité / Souveraineté »	
homme	femme	homme	femme	homme	femme	homme	femme	homme	femme
sauvage	sensuel	punir	danger	puissance	puissance	dieu	dieu	naissance	mariage
aventurier	rêver	danger	rompre	richesse	richesse	foi	foi	méfiance	naissance
sensuel	voluptueux	mort	mort	or	or	âme	âme	faute	famille
rêver	nudité	rompre	cri	argent	argent	prêtre	sacre	doute	cérémonie
orage	aventurier	angoisse	orage	gloire	gloire	sacre	prêtre	rigide	adorer
nudité	original	labyrinthe	désert	conquérir	conquérir	recueillem.	recueillem.	caresse	doute
original	légèreté	doute	critiquer	vitesse	ruse	cérémonie	noble	masque	enfance
île	bohème	cri	révolte	fusil	ambition	éternel	créateur	tendresse	éternel
révolte	désir	vide	punir	commander	séduire	bijou	souverain	angoisse	dieu
océan	île	critiquer	vide	attaquer	bijou	adorer	absolu	cri	foi
loi	sacrifice	caresse	acheter	charitable	pardon	utilitaire	confiance	volupté	inventeur
dieu	règle	richesse	hériter	arbre	charitable	réfléchir	ruse	art	bâtitteur
foi	prêtre	argent	maison	pardon	paix	produire	gaieté	prêtre	produire
prêtre	économiser	gratuit	récompense	écrire	poésie	robuste	volontaire	recueille	astucieux
économiser	industrie	or	argent	enseigner	théâtre	astucieux	honnête	âme	précision
soldat	soldat	gaieté	richesse	théâtre	nid	logique	punir	noble	audace
morale	patrie	bijou	cadeau	école	recueillem	concret	politesse	sacre	maîtriser
obéir	morale	récompense	or	art	art	maîtriser	méfiance	foi	robuste
discipline	obéir	confort	confort	poésie	livre	précision	effort	dieu	concret
patrie	discipline	cadeau	bijou	livre	arbre	efficace	puissance	créateur	efficace

Tableau 2.8 : Confrontation, axe par axe, des analyses réalisées séparément sur deux classes d'âges

axe 2 : « Devoir/ Plaisir »	axe 3 : « Attach. / Détach. »	axe 3 : « Attach. / Détach. »	axe 2 : « Devoir/ Plaisir »	axe 4 : « Esprit / Matière »		axe 5 : « Cœur / Raison »		axe 6 : « Humilité / Souveraineté »	
				- 45 ans	+ 45 ans	- 45 ans	+ 45 ans	- 45 ans	+ 45 ans
orage	révolte	fleur	cadeau	richesse	richesse	dieu	dieu	faute	naissance
sauvage	orage	gaieté	parfum	or	puissance	foi	foi	naissance	doute
étranger	sauvage	douceur	moelleux	vitesse	or	âme	âme	rigide	masque
lune	bohème	tendresse	voluptueux	puissance	argent	prêtre	prêtre	doute	caresse
révolte	désert	paix	bijou	séduire	conquérir	sacre	sacre	méfiance	tendresse
feu	nudité	cadeau	rire	argent	fusil	recueillem.	recueillem.	caresse	cérémonie
bohème	original	maternel	rêver	sensuel	gloire	cérémonie	éternel	nœud	faute
mystère	feu	caresse	gaieté	désir	commander	éternel	absolu	interdire	mariage
désert	aventurier	parfum	sensuel	conquérir	hériter	bijou	infini	tendresse	immobile
différent	ironie	rêver	caresse	gloire	vitesse	souverain	noble	masque	méfiance
économiser	discipline	méfiance	rigide	travail	modestie	construire	sensuel	audace	sublime
hériter	soldat	vide	danger	vieillir	amitié	réfléchir	dynamique	art	astucieux
obéir	mariage	armure	interdire	enseigner	enseigner	robuste	utilitaire	noble	noble
mariage	honneur	rompre	cri	dieu	livre	astucieux	astucieux	âme	robuste
patrie	richesse	muraille	discipline	obéir	paix	utilitaire	produire	sacre	élite
propriété	économiser	angoisse	rompre	prêtre	consoler	logique	matériel	prêtre	audace
gloire	hériter	fusil	angoisse	morale	charitable	maîtriser	volontaire	recueillem.	créateur
richesse	patrie	danger	sacrifice	foi	école	précision	efficace	créateur	efficace
discipline	propriété	punir	obéir	recueillem.	confiance	concret	logique	foi	concret
argent	argent	attaquer	punir	école	pardon	efficace	précision	dieu	maîtriser

Les nuances relevées proviennent essentiellement des classements des mots dans l'explication des axes concernés, classements qui peuvent varier légèrement sans perturber leur interprétation initiale. Le tableau 2.7 présente, pour chaque axe<sup>1</sup>, les dix premiers mots qui leur sont le plus fortement corrélés.

Les trois premiers axes de la structure sémiométrique (axes 2-3-4) sont stables. Les valeurs révélées par ces axes sont similaires chez les hommes et chez les femmes et ce malgré des différences notables observées dans l'analyse commune, notamment pour le troisième facteur. Aussi la notion d'« Attachement » de l'axe 3 est une valeur plus féminine qui existe cependant chez les hommes et celle de « Détachement » est une valeur plus masculine que l'on retrouve aussi chez les femmes. On note pour le pôle « Attachement », chez les femmes, la présence des mots *Acheter*, *Hériter*, *Fidélité* et l'absence de *Gaieté*, *Caresse*, *Rire*, *Douceur* dans la liste des dix premiers mots. Il existe donc des nuances dans la caractérisation de l'axe, nuances qui s'estompent ou évoluent si l'on retient plus de mots pour caractériser cette extrémité de l'axe.

C'est sans doute pour les axes 5 et 6 que se produisent quelques irrégularités dans la mesure où les quatre concepts émergeant de ces axes dans l'analyse globale ne correspondent pas exactement aux axes de même rang dans l'analyse des hommes et dans celle des femmes. Mais, hormis ce décalage sur les deux derniers axes qui s'estomperait sans doute avec un échantillon plus grand, la structure sémiométrique observée sur ces deux sous-populations est stable, stabilité que l'on retrouve d'ailleurs dans l'analyse suivante, celle par âge.

#### **2.1.4 Stabilité selon l'âge**

Nous venons de voir que les axes 3 « Attachement / Détachement » et 5 « Raison / Cœur » étaient liés, dans l'analyse globale, au sexe des répondants, mais qu'ils subsistaient lorsqu'on élimine l'effet *sexe*, c'est-à-dire en travaillant séparément sur deux sous-populations de même sexe. Or l'axe 2 (« Devoir / Plaisir ») et, dans une moindre mesure, l'axe 4 (« Esprit / Matière ») sont, eux, liés à la variable « âge ». Que se passe-t-il si l'on scinde la population en deux catégories d'âge : les « moins de 45 ans » et les « plus de 45 ans », et si l'on procède aux analyses à l'intérieur de chacune de ces deux sous-populations ?

---

1. Le premier facteur dont le statut est particulier n'est pas pris en compte ici. Toutefois, il se révèle être également stable.

Là encore, on observe une permanence des structures à l'exception d'une simple interversion des axes 2 et 3 pour les personnes de plus de 45 ans, mettant alors en second plan le « Plaisir / Devoir » par rapport à l'« Attachement / Détachement » (cf. tableau 2.8)<sup>1</sup>.

Ainsi, la notion de « Plaisir », chez les « moins de 45 ans » se rapproche, du point de vue des mots qui la caractérise, de la notion d'« Esprit » (*Lune, Mystère, Art, Inconnu, Fleuve, Océan*). De la même façon, la notion de « Devoir » pour cette même sous-population, fait plus référence à la notion de discipline et à une certaine réussite matérielle, proche du pôle « Matière » de l'axe 4.

L'« Attachement » de l'axe 3 est moins matériel chez les « moins de 45 ans » (absence, parmi les 15 premiers mots, des mots : *Or, Récompense, Argent*) et plus lié aux mots du « Plaisir » chez les « plus de 45 ans » (*Voluptueux, Sensuel, Séduire*). Le « Détachement » est plus agressif et guerrier pour les plus jeunes (*Attaquer, Fusil, Armure, Guerre*) et plus oblatif pour les plus âgés (*Obéir, Sacrifice, Interdire, Effort*).

La notion d'« Esprit » de l'axe 4 est moins bien définie pour les moins de 45 ans dont les valeurs se retrouvent en partie dans le « Plaisir » de l'axe 2.

Précisons que pour ces études de stabilité selon l'âge, comme pour celles relatives au sexe dans la section précédente, la division en deux de l'échantillon déséquilibre fortement celui-ci, de façon évidente pour les critères incriminés, mais aussi pour d'autres critères (profession, niveau d'éducation, région...) qui n'étaient pas forcément équilibrés à l'intérieur de chaque tranche d'âge ou de chaque sexe. Les sous-échantillons n'étant plus représentatifs, la permanence globale des structures n'en est que plus surprenante, au contraire des petites différences observées qui paraissent finalement modestes.

## 2.2 Intervalles de confiance des valeurs propres

Dans cette section, plus technique, ainsi que dans les sections suivantes, la stabilité et la cohérence de la structure sémiométrique sera éprouvée en utilisant les outils de la validation statistique et de la simulation. On est naturellement conduit à poser un certain nombre de questions sur la qualité des représentations : Observe-t-on vraiment quelque chose ? Les données ont-elle une structure ? Ou, au contraire, de simples fluctuations

---

1. On ne s'étonnera pas, compte tenu de l'intervention d'axes signalée, de voir sur le tableau 2.8 l'axe 2 des « moins de 45 ans » comparé à l'axe 3 des « plus de 45 ans » (deux premières colonnes) et réciproquement (deux colonnes suivantes).

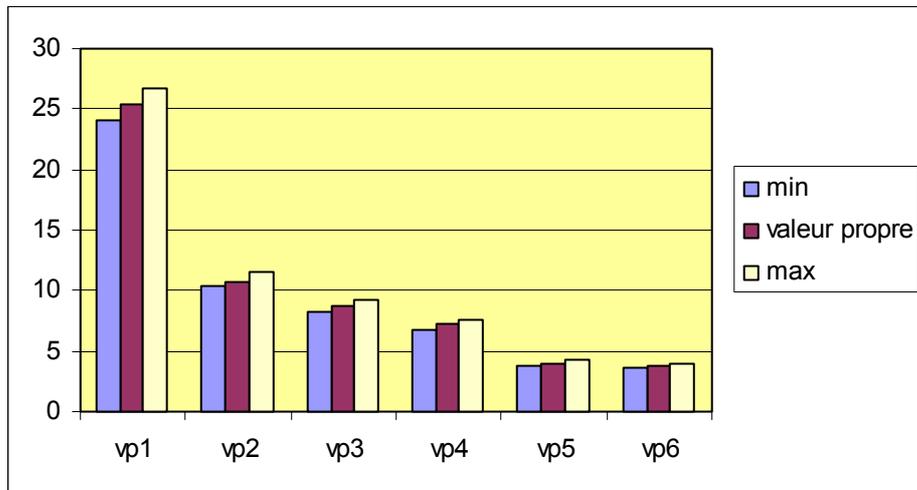
d'échantillonnage suffiraient-elles à expliquer les pourcentages de variance obtenus (pourcentages qui, rappelons-le, mesurent l'importance des axes) ?

Cette question sera tout d'abord traitée dans un cadre statistique classique : il existe en effet une procédure permettant d'évaluer la confiance que l'on peut accorder aux premières *valeurs propres*, c'est-à-dire finalement aux premiers axes. L'ampleur de l'intervalle de confiance qui va être calculé donne une indication sur la stabilité de la valeur propre vis-à-vis des fluctuations dues à l'échantillonnage. L'intervalle de confiance d'Anderson (*cf.* annexe A1.9.4) pour chaque valeur propre décrit en quelque sorte l'incertitude sur celle-ci. Le tableau 2.9, puis la figure 2.6 dans le cas des deux premières valeurs propres, présentent les intervalles de confiances pour les 6 premières valeurs propres dans le cas de trois populations de tailles différentes.

**Tableau 2.9 : Intervalles de confiance des 6 premières valeurs propres (cas de trois tailles d'échantillon différentes)**

		borne inférieure	valeur propre	borne supérieure
Echantillon 2 000	vp1	24.00	25.35	26.77
	vp2	10.40	10.98	11.60
	vp3	8.24	8.70	9.19
	vp4	6.80	7.18	7.58
	vp5	3.80	4.01	4.23
	vp6	3.59	3.79	4.00
Echantillon 10 000	vp1	25.49	26.19	26.91
	vp2	10.07	10.35	10.63
	vp3	8.58	8.82	9.06
	vp4	6.82	7.01	7.20
	vp5	4.04	4.15	4.26
	vp6	3.58	3.68	3.78
Echantillon 15 000	vp1	26.40	26.99	27.59
	vp2	10.13	10.36	10.59
	vp3	8.64	8.84	9.03
	vp4	6.78	6.93	7.09
	vp5	4.07	4.16	4.26
	vp6	3.61	3.69	3.77

Les calculs portent sur trois échantillons sémiométriques de 2 000, 10 000 et 15 000 individus. Les résultats sont très satisfaisants, en ce sens que, pour une taille d'échantillon donnée, les intervalles de confiance des premières valeurs propres n'empiètent pas, ce qui traduit bien que celles-ci sont significativement distinctes.



**Figure 2.6 : Variations des intervalles de confiance des six premières valeurs propres (vp1,...vp6) pour une taille de l'échantillon de 2000.**

On lit, par exemple, sur les deux premières lignes du tableau 2.9, comme sur la figure 2.6, que la borne supérieure de l'intervalle de la seconde valeur propre (11.60) est plus petit que la borne inférieure de celui de la première (24.00).

La figure 2.6 montre aussi, dans le cas le plus défavorable (échantillon de taille 2000), que les intervalles de confiance restent petits par rapport aux valeurs propres. On note cependant que pour cette taille d'échantillon, les valeurs propres 5 et 6 sont voisines et que leurs intervalles de confiance empiètent.

Tout ceci prouve que les premiers axes principaux occupent des directions significatives, et qu'ils sont bien individualisés.

Ces méthodes s'appliquent en fait à des observations distribuées selon une loi normale, mais leur robustesse leur permet d'être utilisées en dehors de ce cadre. Pour être plus rigoureux et s'affranchir de ces hypothèses, nous allons utiliser des validations par la méthode de *bootstrap non paramétrique*.

Nous verrons donc au paragraphe suivant comment les méthodes de validation par rééchantillonnage s'appliquent à l'étude de la stabilité des formes sans hypothèse au préalable sur la distribution des données.

## 2.3 Stabilité par rééchantillonnage

La permanence des structures observée sur des échantillons indépendants devrait nous satisfaire et nous convaincre que, lorsqu'un ensemble de mots a été choisi, une série d'axes ou de dimensions apparaissent assez systématiquement, pas toujours exactement dans le même ordre, mais avec des caractérisations par les mots très voisines.

On ne sait que peu de choses, cependant, sur la stabilité des positions des mots les uns par rapport aux autres. Leur classement sur les axes a bien montré qu'ils devaient être proches d'une population à une autre. Mais on peut se poser légitimement la question de savoir quelle est la précision de la position d'un mot sur les plans engendrés par deux axes principaux, compte tenu des fluctuations d'échantillonnage de nos recueils de réponses.

Les statisticiens ne pouvaient pas répondre à cette question de façon satisfaisante jusqu'à une période récente au cours de laquelle la puissance des nouveaux moyens de calcul a permis de systématiser les méthodes dites de *rééchantillonnage*.

Parmi ces méthodes, une méthode statistique de validation, désignée sous le nom de *bootstrap*<sup>1</sup> permet de vérifier à la fois la stabilité des axes (ce qui ne fera que confirmer les résultats des sections précédentes de ce chapitre), et d'évaluer la confiance que l'on peut accorder à la position de chaque point dans les plans principaux.

La méthode de *bootstrap* classique utilisée ici consiste à perturber l'échantillon initial en simulant un grand nombre d'échantillons de même taille<sup>2</sup>. Dans ce contexte, les simulations sont appelées *réplications*.

Deux variantes de cette méthode seront utilisées, le *bootstrap* partiel<sup>3</sup>, et le *bootstrap* total<sup>4</sup>.

Elles donneront des résultats convergents, attestant la stabilité des structures observées jusqu'au sixième axe.

---

1. Cette terminologie imagée due à Efron (1979) se réfère à l'expression « *to pull oneself up by one's own bootstraps* » qui signifie littéralement « s'élever en tirant ses propres lanières de bottes », ou encore « progresser sans demander d'aide » (cf. annexe A1.9.5).

2. Ces échantillons sont obtenus par  $n$  tirages au hasard *avec remise* des individus dans l'échantillon initial de  $n$  individus.

3. Le *bootstrap* partiel consiste à projeter les colonnes des tableaux répliqués (mots) comme des éléments supplémentaires sur les axes de l'analyse de référence, c'est-à-dire l'analyse de l'échantillon initial non perturbé. Avec 30 réplications, on obtient 30 positions (permettant de calculer des ellipses de densité) pour chaque mot (cf. annexe A1.9.5).

4. Le *bootstrap* total consiste à refaire des analyses en composantes principales complètes sur chaque échantillon répliqué. Comme dans le cas partiel, avec 30 réplications, on obtient 30 positions (permettant de calculer des ellipses de densité) pour chaque mot (cf. annexe A1.9.5).

### 2.3.1 Bootstrap partiel

Les figures 2.6 et 2.7 montrent quelques ellipses de confiance obtenues par la méthode de validation selon les principes du *bootstrap* partiel respectivement pour les plans (2, 3) et (5, 6). Le plan (2, 3) est le premier plan sémiométrique<sup>1</sup>, le plus utilisé dans les visualisations, et le plan (5, 6) est le dernier plan sémiométrique.

Il s'agit d'un échantillon « France », regroupant les quatre échantillons indépendants des années 1990, 1996, 1998 et 1999, et totalisant 11055 individus, représentatifs de la population française âgée de 18 ans et plus. Les ellipses sont de tailles très réduites, et donc les fluctuations de la position des points sont très faibles, malgré la sévérité du principe des réplifications.

En effectuant un tirage avec remise dans l'échantillon, on montre qu'environ 30% des individus sont exclus de chaque réplification ; d'autres individus en revanche apparaissent deux, trois fois, jusqu'à six fois dans certains cas exceptionnels.

On note que les ellipses relatives au plan (5, 6) sont légèrement plus grandes, tout en laissant à la position des points une précision acceptable.

### 2.3.2 Bootstrap total

Ce sont les figures 2.8 et 2.9 qui montrent les mêmes plans principaux que précédemment, avec maintenant des ellipses plus grandes<sup>2</sup>.

L'examen de la figure 2.9 montre que les zones de confiance des mots interfèrent peu. Il s'agit pourtant d'une épreuve sévère (le *bootstrap* total) et d'un plan sémiométrique *a priori* moins stable que le plan principal, le plan (5, 6) ; les effets des fluctuations d'échantillonnage sont donc faibles, confirmant ainsi la stabilité de ces facteurs et des structures obtenues.

---

1. Les axes 2 et 3 sont les deux axes les plus importants, puisque l'axe 1 ne fait pas partie de la structure sémiométrique proprement dite.

2. En fait, le *bootstrap* total donne une vue pessimiste de la précision des points, car les analyses en composantes principales sont refaites pour chaque réplification, et donc les plans de projection sont calculés sur des données perturbées : non seulement les données sont perturbées, mais l'outil de visualisation est lui-même perturbé (cf. annexe A1.9.5).



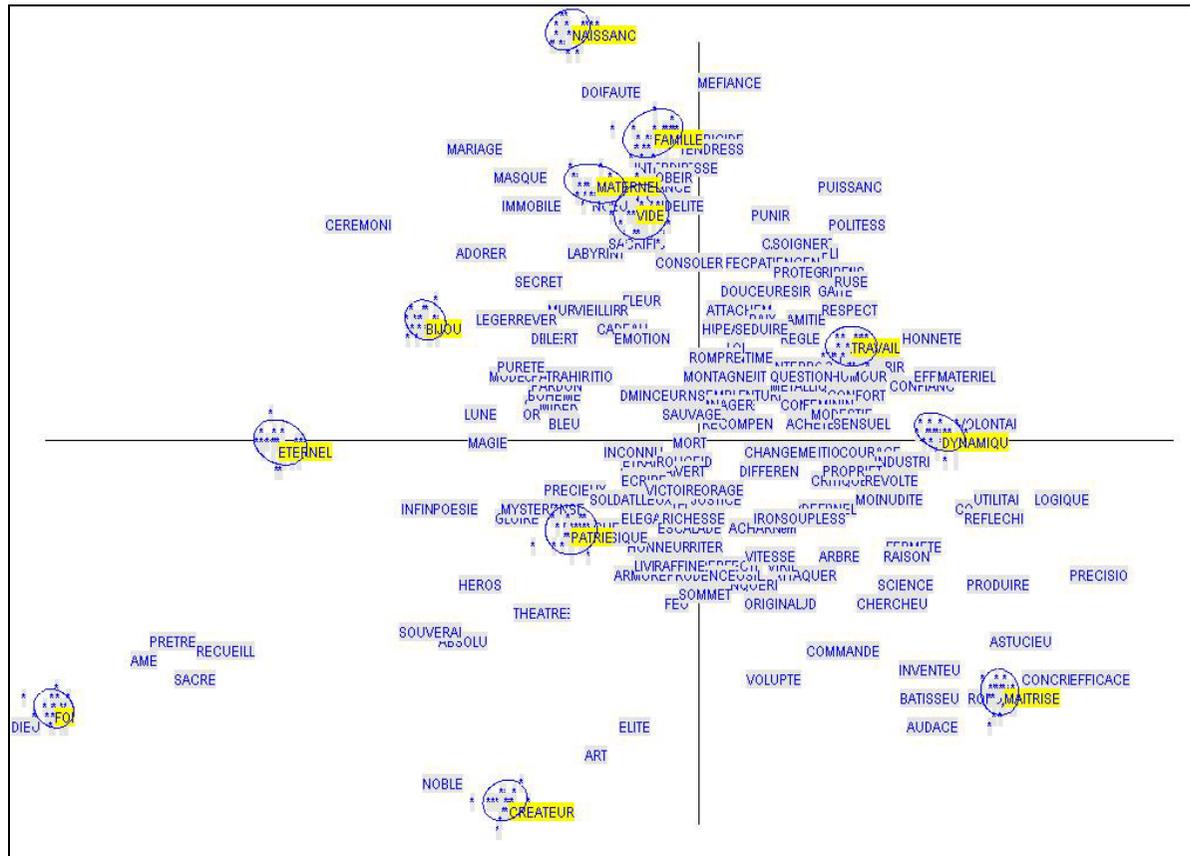
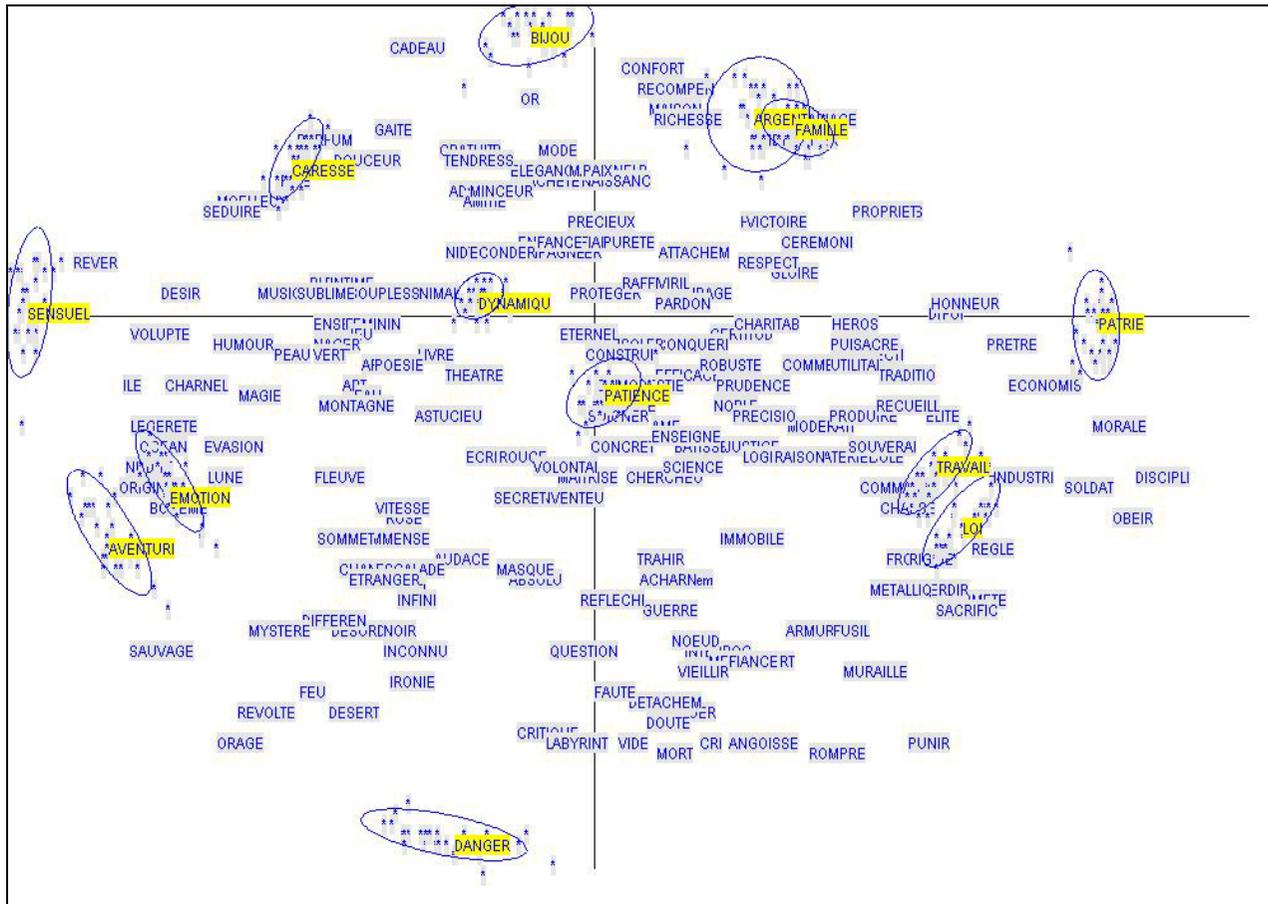


Figure 2.7 : Plan engendré par les axes 5 et 6 – Quelques zones de confiances (*Bootstrap partiel*)

## Stabilité de la structure sémiométrique

Figure 2.8 : Plan engendré par les axes 2 et 3 – Quelques zones de confiance (*Bootstrap total*)

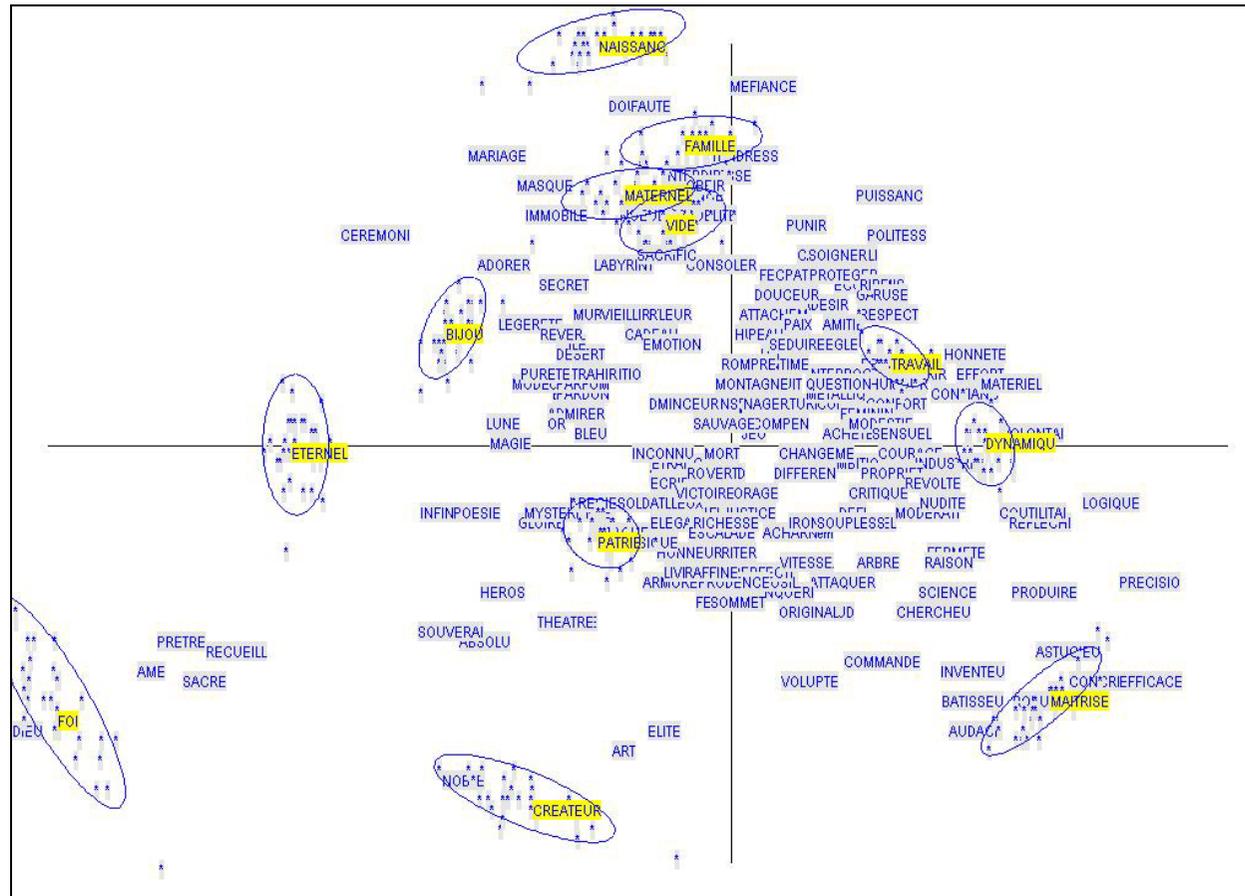


Figure 2.9 : Plan engendré par les axes 5 et 6 – Quelques zones de confiance (*Bootstrap total*)

## 2.4 Stabilité structurelle

Dans la section précédente de ce chapitre, c'est la stabilité des structures vis-à-vis des fluctuations d'échantillonnage qui a été éprouvée sur l'ensemble des individus selon la méthode de *bootstrap*.

Nous allons changer de registre et poser une question radicalement différente, et beaucoup plus difficile : dans quelle mesure les structures sémiométriques dépendent-elles du choix des mots ?

Cette question est effectivement plus difficile car il n'existe pas un univers des mots comparable à l'univers des individus susceptibles d'être interrogés dans une enquête, univers constitué, par exemple, par l'ensemble des individus âgés de 18 ans ou plus d'un pays donné. Les statisticiens ont l'habitude de traiter la dimension « individu » ou « observation » à laquelle on peut associer la notion de tirage dans une urne et les modèles mathématiques correspondant. Rares sont les situations où l'on est amené à parler de tirage ou d'univers de variables<sup>1</sup>.

On va néanmoins chercher à tester la stabilité des structures vis-à-vis de l'ensemble des variables. On fait alors l'hypothèse que les variables (les 210 mots) sont choisies par un tirage aléatoire sans remise dans une urne contenant l'ensemble des mots « sémiométrisables » de la langue française<sup>2</sup>. Un tirage dans cet univers suppose à chaque fois la constitution d'un nouveau questionnaire et un nouveau travail de terrain, opérations dont le coût serait prohibitif. Comme dans le cas des échantillons d'individus, le *bootstrap* fournit une solution pragmatique et économique à ce problème : on va construire des répliques de l'ensemble des mots en procédant à des tirages avec remise dans l'ensemble des 210 mots déjà sélectionnés.

Alors que le *bootstrap* classique réalisé sur les individus teste la stabilité des *patterns* de variables, c'est-à-dire de mots (*cf.* section 2.2), le *bootstrap* réalisé sur les variables permet d'éprouver les *patterns* observables au niveau des individus. Or ceux-ci sont anonymes et nombreux (plusieurs milliers). Les seules structures intéressantes sont celles qu'ils induisent sur leurs caractéristiques de base (sexe, âge, profession, niveau d'éducation, activités, types d'habitats...). La procédure décrite en annexe A1.2.4 permet

---

1. Il est des circonstances où le problème de l'échantillonnage se pose à la fois pour les variables et les individus. Ce problème ancien, posé par Hotelling (1933) [inventeur/découvreur de l'analyse en composantes principales], a été en particulier étudié par Escoufier (1970).

2. Il s'agit, rappelons-le, de mots pleins - pas de mots outils ou grammaticaux - faisant partie du vocabulaire de base, non consensuels, non ambigus sémantiquement - non polysémiques -, chargés émotionnellement ou axiologiquement - pas de mots neutres du type *table*, ni de mots techniques comme *magnétoscope*.

cependant de positionner les mots correspondant à chaque échantillon répliqué, y compris les mots qui sont absents de certaines répliques<sup>1</sup>.

L'ensemble des variables (ici 210 mots) est beaucoup plus restreint que celui des individus (plusieurs milliers). En moyenne, à chaque tirage avec remise, on démontre qu'environ deux tiers des variables initiales seront présentes, ce qui constitue une ponction très sévère dans l'ensemble initial, et donc une perturbation que l'on peut juger excessive du tableau de données.

La figure 2.10 montre les résultats de cette procédure sur le plan des axes [2,3] : les ellipses de confiance des mots sont bien sûr très dilatées par rapport à celles des graphiques précédents, mais les résultats n'ont rien de catastrophique. Le long du second axe (horizontal), la zone de confiance elliptique du mot *Sensuel* reste opposée à celle du mot *Patrie*, de même, le long du troisième axe (vertical) la zone de confiance du mot *Bijou* reste opposée à celle du mot *Danger*. Les interprétations de ces axes résistent aux très fortes perturbations imposées au recueil de mots.

Les axes suivants montrent une stabilité moins nette, mais on peut penser qu'une liste de mots plus étendue (par exemple une liste de 306 mots, ce qui était le cas lors des premières versions de la sémiométrie) permettrait d'aller plus loin dans l'application de cette épreuve de validité par ailleurs extrêmement sévère.

Il faut retenir de cette section que cette procédure (novatrice) de *bootstrap* sur variables fournit une présomption - et non une preuve - de stabilité de la structure vis-à-vis du choix des mots. Le chapitre 4, consacré à la sémiométrie « ouverte » apportera également des contributions à ce problème de la stabilité vis-à-vis de la composition de la liste, problème intéressant et important, mais pas crucial si la sémiométrie n'est utilisée que pour comparer et différencier, ce qui est le cas dans le cadre des applications actuelles.

---

1. En bref : une réplique *bootstrap* consiste en un tirage avec remise au sein des 210 mots de départ. Certains mots seront absents (ils auront un poids nul), d'autres apparaîtront une fois (poids 1), deux fois (poids 2), etc. En remplaçant les poids nuls par des poids infinitésimaux, ce qui nécessite quelques aménagements techniques supplémentaires, on verra se positionner automatiquement les mots absents comme mots supplémentaires (cf. annexe A1.9.3).

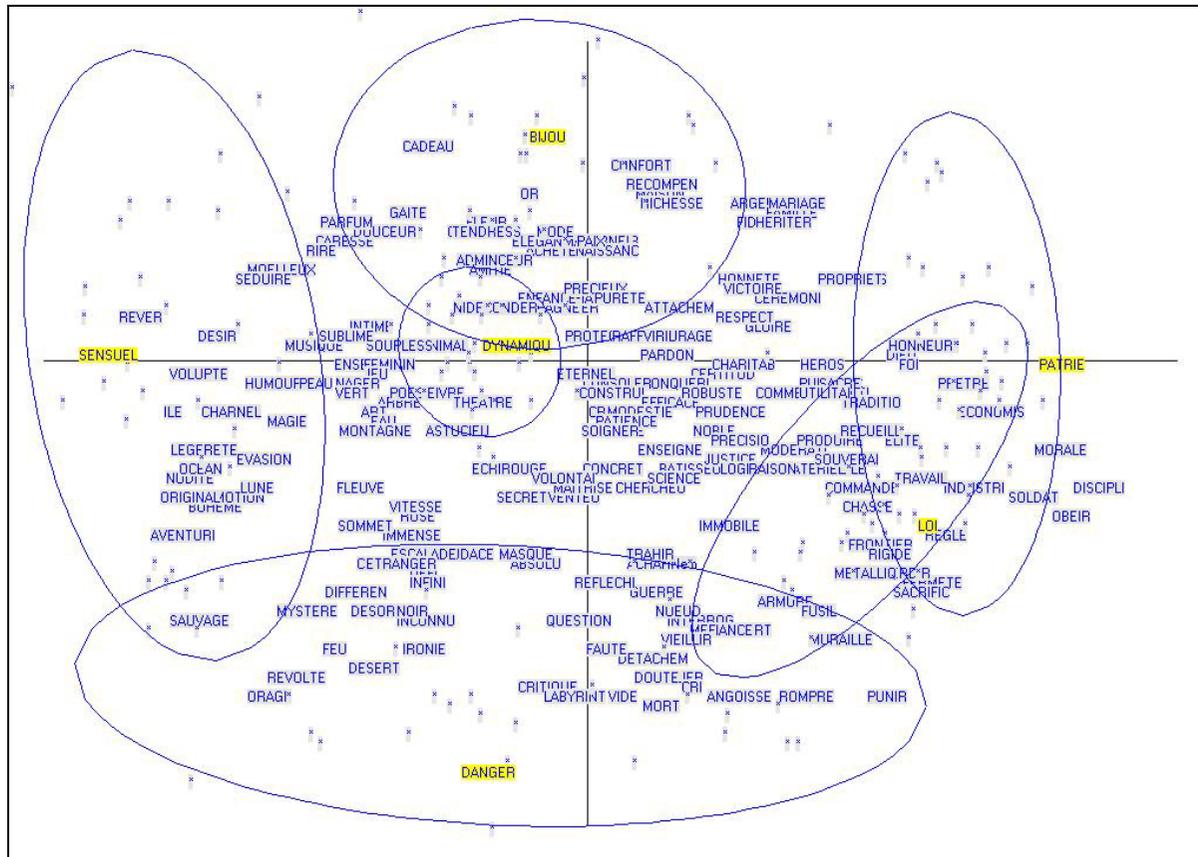


Figure 2.10 : Plan engendré par les axes 2 et 3 – Zones de confiances (*Bootstrap sur variables*), ellipses de confiance des mots : *Bijou, Danger, Dynamique, Loi, Patrie, Sensuel*

## 2.5 Stabilité vis-à-vis du codage

La structure sémiométrique (existence et permanence de six axes principaux caractérisés par les mêmes mots) se révèle stable dans l'espace, dans le temps, et vis-à-vis d'éventuelles fluctuations d'échantillonnage.

On pourrait aussi objecter, assez légitimement, que le codage de base relève d'une convention assez arbitraire : pour le répondant, c'est surtout l'ordre des notes qui importe, et non la matérialisation de celles-ci sous forme d'échelle de 1 à 7, avec des notes en progression arithmétique.

Est-ce que la structure sémiométrique mise en évidence serait la même si le codage proposé était différent, en respectant toutefois l'ordre des notes ?

Notons déjà que toute transformation  $y$  du codage initial  $x$  de la forme  $y = ax + b$  laisse invariante les corrélations, et donc laisse identique toute la structure sémiométrique. Ainsi, on ne change pas les résultats en remplaçant (1, 2, 3, 4, 5, 6, 7) par (2, 4, 6, 8, 10, 12, 14) ou encore par (11, 12, 13, 14, 15, 16, 17).

Pour déformer encore plus le codage initial, il sera donc procédé à des transformations *non-linéaires*<sup>1</sup> (voir figure 2.11). On va voir que la structure est invariante pour une large famille de telles transformations.

Le résultat est le suivant : pour toutes ces transformations, les six premiers axes sémiométriques sont conservés. Ceci signifie que les axes apparaissent dans le même ordre et sont caractérisés par les mêmes mots.

---

1. Les transformations non-linéaires suivantes du codage arithmétique de départ ont été réalisées (voir figure 2.11) : a)  $y = \sqrt{x}$  ; b)  $y = x^2$  ; c)  $y = \log_{10}(1+x)$  ; d) codage « ternaire » selon lequel  $y = 1$  pour  $x = 1$  ou 2 ;  $y = 2$  pour  $x = 3$  ou 4 ou 5 ;  $y = 3$  pour  $x = 6$  ou 7.

Sur la figure 2.11, le carré a été divisé par 5 et le logarithme multiplié par 10 pour ramener les différents codages à une échelle comparable à celle du codage initial (ou codage de base, en pointillé).

Partant un échantillon de 5 527 individus (France), et donc d'un tableau de notes (5 527 x 210), quatre nouveaux tableaux de mêmes dimensions ont été calculés en appliquant aux notes les transformations précédentes.

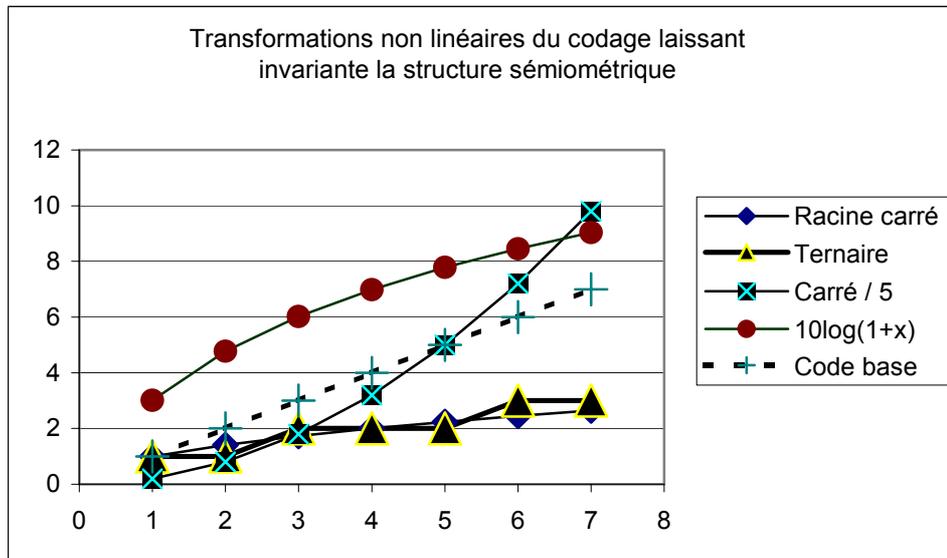


Figure 2.11 : Diverses transformations de l'échelle des notes

Tableau 2.10 : Valeurs propres des différents tableaux recodés

Numéro	Base	Carré	Log	Ternaire	Racine
1	26.6	28.7	23.5	22.1	25.1
2	10.2	10.3	9.6	8.2	9.9
3	8.7	8.4	8.7	7.1	8.8
4	7.0	7.0	6.4	5.9	6.8
5	4.1	4.0	4.0	3.6	4.1
6	3.6	3.7	3.3	3.2	3.5

*On constate que la première valeur propre est maximale pour les carrés des notes (codes fortement dilatés, de 1 à 49), est minimale pour le codage ternaire, qui contracte les 7 valeurs en trois valeurs.*

Ce sont les directions propres (ou axes principaux) qui sont stables, c'est-à-dire caractérisées par les mêmes groupes de mots. Les valeurs propres, qui décrivent les allongements le long de ces axes, sont, elles, plus sensibles au codage, sans pour autant modifier l'ordre des axes.

Il est intéressant de noter que le codage ternaire n'altère pas la structure : cela indique que l'échelle en 7 positions est probablement assez redondante. Mais cela ne prouve pas obligatoirement (il faudrait le vérifier, mais cela impliquerait de retourner sur le terrain) qu'une échelle en trois positions, proposée directement aux personnes interrogées, restituerait la structure initiale.

En revanche, un codage binaire, (par exemple :  $y = 1$  pour  $x = 1,2,3,4$  ;  $y = 2$  pour  $x = 5,6,7$ ) trop destructeur d'information, altère sensiblement la structure. Le premier axe est à peu près conservé (valeur propre 19.3,

inférieure à toutes les premières valeurs propres du tableau 2.4.1) mais on observe une rotation à l'intérieur du plan (2,3). De façon surprenante, les axes 4, 5, 6, sont cependant conservés. Avec un tel codage, certains mots comme *guerre*, *mort*, *trahir*, n'obtiennent que des notes 1 sur tous les individus, d'autres, comme *paix* ou *tendresse* n'obtiennent que des notes 2.

De tels mots, dont les variances sont nulles, n'interviennent plus dans les corrélations et sont exclus de la structure.

## 2.6 Analyse en facteurs communs et spécifiques

La méthode d'analyse en facteurs communs et spécifiques, qui remonte au tout début du vingtième siècle (*cf.* annexe A1.6), est apparentée à l'analyse en composantes principales, mais intervient en général dans un contexte fort différent : il ne s'agit plus d'un simple instrument d'observation, mais d'un modèle.

Ce modèle stipule ici que chacune des 210 notes (correspondant donc à chacun des mots) dépend d'un petit nombre de facteurs communs et d'un facteur spécifique (voir la formulation exacte en annexe A1.6).

Les facteurs communs correspondraient, par exemple, aux six premières dimensions<sup>1</sup>, et les facteurs spécifiques résumeraient en quelque sorte les 204 axes suivants, qui ne sont pas pris en compte.

Ce modèle analytique, qui exige d'ailleurs quelques contraintes supplémentaires, a bien peu de chance d'être validé par un test statistique rigoureux, même en prenant une trentaine ou une cinquantaine de facteurs communs, ne serait-ce que parce que les notes de départ sont des variables discrètes et non des variables continues. Et pourtant, l'ajustement d'un tel modèle (en utilisant la procédure évoquée dans l'annexe A1.6) avec, par exemple, 12 facteurs communs, livre pour les six premiers facteurs (en fait les coefficients des mots sur ces facteurs, appelés  *saturations* ) les six premiers axes de la structure sémiométrique.

L'application de cette méthode figure dans ce chapitre dévolu à la stabilité des résultats car la quasi-totalité des conclusions précédentes sur l'interprétation des axes s'applique aussi à l'analyse factorielle en facteurs communs et spécifiques. Les résultats ont donc résisté à des perturbations importantes des données de base, à des changements de codage, et également à des modifications notables de la méthode de réduction des données.

---

1. L'analyse en composantes principales est dans ce cas un cas particulier d'analyse en facteurs communs et spécifiques avec 210 facteurs communs (on garde tous les axes) et aucun facteur spécifique. Le facteur spécifique est présent pour modéliser les axes abandonnés.

Un des sous-produits intéressants de ce dernier type d'analyse est le calcul, pour chaque mot, de sa variance spécifique (variance non expliquée par les facteurs communs) dont le complément à 1 est la variance commune (appelée parfois *communauté*).

**Tableau 2.11 : Mots ayant les variances communes les plus fortes... et les plus faibles**

Communautés les plus fortes			Communautés les plus faibles		
Mot	Spécificité	Commun.	Mot	Spécificité	Commun.
foi	0.332	0.668	justice	0.823	0.177
dieu	0.362	0.638	chasse	0.825	0.175
prêtre	0.447	0.553	frontière	0.825	0.175
recueillement	0.480	0.520	détachement	0.828	0.172
sacré	0.491	0.509	nager	0.829	0.171
puissance	0.525	0.475	interdire	0.832	0.168
discipline	0.538	0.462	métallique	0.836	0.164
efficace	0.538	0.462	changement	0.839	0.161
sensuel	0.559	0.441	jeu	0.849	0.151
richesse	0.562	0.438	inconnu	0.853	0.147
argent	0.568	0.432	nœud	0.854	0.146
gloire	0.571	0.429	évasion	0.856	0.144
or	0.585	0.415	noir	0.866	0.134
poésie	0.587	0.413	acharnement	0.870	0.130
séduire	0.588	0.412	guerre	0.891	0.109
arbre	0.592	0.408	trahir	0.893	0.107
âme	0.593	0.407	immobile	0.914	0.086
élégance	0.595	0.405	rouge	0.925	0.075

Le tableau 2.11 donne la liste des 18 mots dont les notes sont les mieux expliquées par les facteurs communs (mots ayant les *communautés* les plus fortes) et des 18 mots dont les notes sont les moins bien expliquées par les facteurs communs (mots ayant les spécificités les plus fortes)<sup>1</sup>.

1. Un résultat assez similaire pourrait être trouvé en sommant, pour les six premiers axes, les carrés des coordonnées des mots sur les axes, qui sont d'ailleurs les corrélations des mots avec ces axes (tableau A2.3 de l'annexe 2). Ces sommes représentent, très approximativement, les communautés, alors que leur complément à 1 représente les spécificités. Il ne s'agit ici que d'une approximation car les vraies communautés de l'analyse factorielle en facteurs communs et spécifiques s'obtiennent par un processus itératif (annexe A1.6).

On retrouve dans la première colonne les mots bien expliqués par l'ensemble des premiers axes de la sémiométrie, et dans la seconde des mots qui sont peu ou mal expliqués par ces premiers axes.

## 2.7 Conclusion

Une fois définie la liste de mots et le protocole d'enquête (le questionnaire et le libellé de la question de base), la structure extraite du tableau de données recueillies apparaît stable. Cette structure, formée de six directions principales ordonnées, ne change pas d'un échantillon à un autre, d'une année à une autre, ne subit pas de modifications (ou subit des modifications minimales) d'un pays à un autre – et ce malgré les risques pris lors des traductions du questionnaire.

La structure sémiométrique se retrouve et se répète à l'intérieur de sous-populations particulières (sexe, classes d'âge) suggérant un phénomène fractal, mais il ne s'agit à ce stade que d'une analogie.

Cette stabilité est également confirmée par les épreuves statistiques standards de rééchantillonnage (*bootstrap*) qui confirment les expériences plus empiriques précédentes. Une épreuve non standard (*bootstrap* sur variables) fut une première tentative d'évaluation du degré de dépendance de la structure obtenue vis-à-vis de la composition de la liste de mots.

Le codage numérique choisi pour les réponses peut-être décliné suivant plusieurs fonctions sans altération de la structure, ce qui lève l'hypothèque d'une sensibilité des corrélations au codage somme toute conventionnel de départ.

Enfin, l'analyse factorielle classique, ou analyse en facteurs communs et spécifiques, reproduit les mêmes premières dimensions, comme d'ailleurs l'analyse logarithmique évoquée plus loin au chapitre 5. Il ne s'agit plus exactement de stabilité, mais de présomption de robustesse de la structure sémiométrique.

Les chapitres 3 et 4 qui vont suivre vont relater plusieurs expériences apportant d'autres éléments de réponse à la question beaucoup plus complexe et ambitieuse de l'éventuel caractère intrinsèque de la structure.